Rainer Teubner

ENTWURF EINES BENUTZERFREUNDLICHEN, DIREKTSPEICHERNDEN INFORMATIONSSYSTEMS FÜR DEN PRÄPARATIV ARBEITENDEN ORGANISCHEN CHEMIKER

Gesellschaft für Information und Dokumentation
Sektion für Systementwicklung
Heidelberg

Diss. Frankfurt a/M.
1973

UNIV.
LABRADOR
LIBR.

Entwurf eines benutzerfreundlichen, direktspeichernden Informationssystems für den präparativ arbeitenden organischen Chemiker

INAUGURAL - DISSERTATION
zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt bei dem Fachbereich Chemie der
Johann-Wolfgang-Goethe-Universität
zu Frankfurt am Main

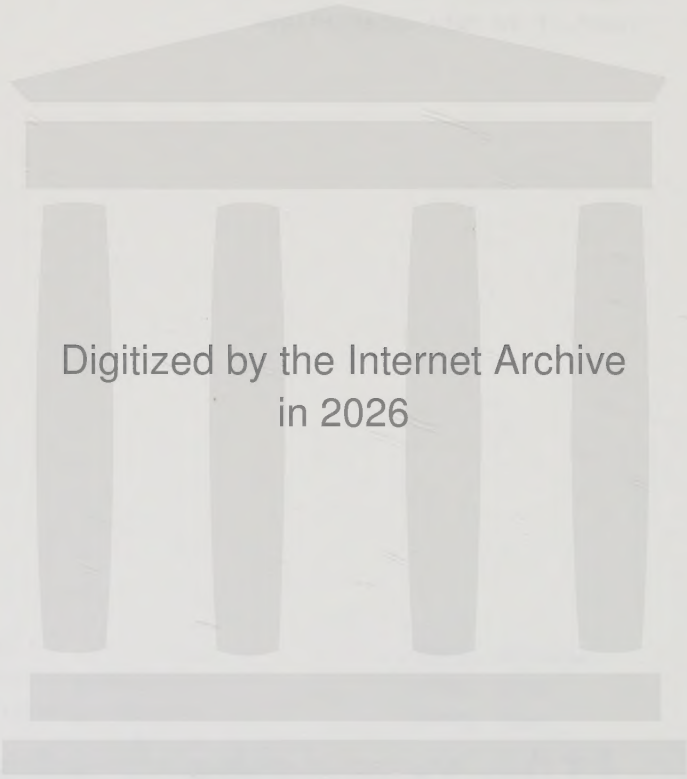
von
RAINER TEUBNER
aus
Seligenstadt

Seligenstadt
1983

Gedruckt mit Genehmigung des Fachbereichs Chemie
der J.W. Goethe-Universität in Frankfurt am Main

Dekan: Prof. Schleip
1.Berichterstatter: Prof. Ried
2.Berichterstatter: Prof. Kunz
Tag der Promotion: 10.11.83

Diese Arbeit wurde in der Zeit von Herbst 1977 bis Herbst 1982 in der Arbeitsgruppe für Informationswissenschaft in der Chemie an der Johann-Wolfgang-Goethe-Universität zu Frankfurt am Main angefertigt.


Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41523704_0002

Meinen verehrten Lehrern

Herrn Prof.Dr.W.Kunz

und

Herrn Prof.Dr.W.Ried

danke ich herzlich für die
ständige Hilfsbereitschaft,
für die vielfältigen An-
regungen und das große
Interesse, das sie dieser
Arbeit entgegengebracht
haben.

... Eine Einleitung ist freilich notwendig. Nun, sehen Sie, die ganze Informationstheorie ist in den Köpfen einiger weniger Leute entsprungen, fast über Nacht, anfangs erschien alles verhältnismäßig einfach - Rückkopplung, Homöostase, Information als Umkehrung der Entropie - aber bald hatte sich herausgestellt, das sich das nicht so schnell zum System zusammenschließen läßt, daß es Sumpfland ist, mathematisches Sumpfland, Weglosigkeit. Schulen begannen zu entstehen, die Praxis ging ihren eigenen Weg, baute diese verschiedenen Maschinen, zum Rechnen, zum Übersetzen, lernende, schachspielende... und die Theorie ging auch ihren Weg, so daß sich ein Ingenieur für solche Maschinen mit einem Informationstheoretiker bald nur noch schwer verständigen konnte...

Stanisław Lem: Erzählungen
Formuła Lymphatera
Frankfurt 1979

Inhaltsverzeichnis

1.	Einleitung und Erläuterung der Problematik.....	1
1.1.	Information und Informationswissenschaft.....	2
1.2.	Dokumentation.....	9
1.3.	Kommunikation.....	11
1.4.	Forschung.....	14
1.5.	Informationssysteme.....	18
1.5.1.	Die Handbücher.....	18
1.5.2.	Literaturinformationssysteme.....	19
1.5.3.	Maschinenlesbare Dienste..	22
1.5.4.	Fakten- und Dateninformationssysteme.....	25
1.5.5.	Referral-Systeme.....	25
1.6.	Spezielle Chemie-Informationssysteme.....	26
2.	Syntheseplanungssysteme...	29
2.1.	Auf einer Stoff- und Reaktionsbibliothek beruhende Systeme.....	33
2.2.	"Voraussetzungslose" Systeme.....	38

3.	Aufgabenstellung und Ziele der Arbeit.....	44
3.1.	Die Direktspeicherung von Sachverhalten.....	44
3.2.	Vorstellung der allgemeinen Konzeption von CHEMIS.....	45
4.	Beschreibung der Systementwicklung.....	47
4.1.	Vorarbeiten.....	47
4.2.	Zur Auswahl des behandelten Gebiets.....	49
4.3.	Systemanalyse des Forschungsprozesses präparativ arbeitender Chemiker.....	53
4.4.	Daten.....	75
4.4.1.	Vorbemerkung.....	75
4.4.2.	Überlegungen zur Datenstruktur in CHEMIS.....	77
4.4.2.1.	Daten zu Verbindungen....	82
4.4.2.2.	Infrarot-Spektren.....	86
4.4.2.3.	Kernresonanz-Spektren....	99
4.4.3.	Reaktionsgleichungen.....	108

5.	Heuristiken.....	117
5.1.	Die Reaktionen der Chlorformamidine.....	124
5.2.	Diskussion der Hydrolyseempfindlichkeit der Chlorformamidine.....	142
6.	Experimenteller Teil.....	173
6.1.	Der benutzte Computer.....	173
6.2.	Abfrage.....	175
6.2.1.	Eigenschaften der Verbindungen.....	179
6.2.2.	Spektraldaten.....	198
6.2.3.	Reaktionsgleichungen.....	214
6.3.	Eingabe	226
6.3.1.	Eigenschaften der Verbindungen.....	228
6.3.2.	Spektraldaten.....	235
6.3.3.	Reaktionsgleichungen.....	246
7.	Zusammenfassung.....	250
8.	Anhang.....	252
9.	Literaturverzeichnis.....	257

1. Einleitung und Erläuterung der Problematik

In der Einleitung wird versucht, die terminologische Vielfalt, die sich im Laufe der Entwicklung um die Begriffe
- Information
- Dokumentation und
- Kommunikation

gebildet hat, darzustellen und in ihrer Bedeutungsvielfalt gegeneinander abzugrenzen. Es wird gezeigt, daß sich die Begriffsinhalte letztendlich alle mit der Generierung, der Übermittlung, der Verarbeitung und der Speicherung von Wissen zum Zwecke des Erkenntnisfortschritts befassen. Die Terminologievielfalt ist daher nicht verwunderlich, eröffnen sich doch mannigfaltige Einstiegsmöglichkeiten in das Gesamtgebiet, je nach Fachrichtung, aus der man sich dem Gebiet der Informationswissenschaft nähert. Führt man eine einheitliche Terminologie ein, so erreicht man neben der durchaus wünschenswerten Präzisierung der Begriffe auch einen Verlust der Vielseitigkeit.

R. CAPURRO (1) stellt die historisch-etymologische Entwicklung des Begriffs "Information" in seiner ganzen Breite dar.

Für die drei Begriffe "Information", "Dokumentation" und "Kommunikation" wird in knapper Form Ähnliches versucht und danach werden die sinnvollen Beschränkungen und notwendigen Erweiterungen vorgenommen.

Mit Hilfe der präzisierten Begriffe soll ihre Bedeutung erläutert werden, die sie im Popper'schen Modell des Forschungsprozesses haben.

Daraus lassen sich Grundanforderungen an Informationssysteme ableiten. Als Vorgaben für den weiteren Ausbau des Systems werden die Ergebnisse der Systemanalyse von Problemlösungs-

prozessen vorgesehener Benutzer verwendet (27, 136). Dem Systementwickler bleibt die Aufgabe, die erhaltenen Ergebnisse auf ihre Verallgemeinerungsfähigkeit und Übertragbarkeit zu untersuchen.

Bei der Auswahl des Fachgebiets wurde pragmatisch vorgegangen: Aufbauend auf den Vorarbeiten von B. RAMI (2) soll versucht werden, die dort entwickelte Vorgehensweise zur Spezifikation und Strukturierung eines Syntheseplanungsunterstützungssystems, die auf dem Gebiet der Aminosäuren und Dipeptide erprobt wurde, auf das weniger übersichtliche Gebiet der Chlorformamidine zu übertragen. Durch die Befragung von Chemikern werden typische Informationsbedürfnisse identifiziert und als Prototypen für die Entwicklung der Direktspeicherungsformate von Sachverhalten herangezogen. Dabei war ein Ziel, die Dateien möglichst benutzerfreundlich aufzubauen und eine Abfragesprache für das auf EDV-Anlagen gespeicherte Wissen zu entwickeln, die den Bedürfnissen der Benutzer möglichst gut angepaßt ist.

1.1. Information und Informationswissenschaft

Eine - nicht vollständige - Sammlung der mit dem Substantiv "Information" zusammen auftretenden bedeutungstragenden Verben soll die Vielfältigkeit und den Facettenreichtum dieses Terminus illustrieren:

- Informationen werden produziert (erzeugen, gewinnen, finden und erschließen).
- Informationen werden übermittelt (senden, sprechen, schreiben, drucken, filmen und verteilen).
- Informationen werden aufgenommen (hören, sehen, lesen, empfangen und aufzeichnen).

- Informationen werden verwertet (verstehen, ignorieren, nutzen, weitergeben, verdichten, selektieren und manipulieren).
- Informationen werden verkauft (überlassen, anbieten und zurückhalten).
- Informationen werden verarbeitet (digitalisieren, speichern, sortieren, ergänzen, kürzen, bewerten und auswerten).

Das Mengenproblem spielt bei der Diskussion der Information eine wichtige Rolle.
Allein auf dem Chemie-Gebiet erscheinen jährlich rd. 500.000 Veröffentlichungen (Zeitschriftenaufsätze, Monographien, Forschungsberichte, Patente und Dissertationen).

Anzahl der in "Chemical Abstracts" insgesamt nachgewiesenen Publikationen

Jahr	Anzahl
19o7	11.847
191o	6o.o2o
1915	175.622
192o	256.122
1925	379.726
193o	586.o29
1935	889.345
194o	1.2o6.377
1945	1.427.158
195o	1.622.559
1955	2.o37.767
196o	2.613.o69
1965	3.487.793
197o	4.712.125
1975	6.4o2.39o

Tab. 1

Ansätze zur Lösung des Mengenproblems befinden sich in der Entwicklung:

- Publikation auf Mikrofilmen (entweder vollständige Artikel oder teilweise, z.B. umfangreiche Meßdaten),
- Synopsenzeitschriften,
- Depotverfahren für experimentelle Daten,
- direkte Kommunikation über Computernetze (4),
- Einrichtung und Betrieb einer elektronischen Zeitschrift und
- Entwicklung von Speichermedien, die bei geringen Kosten und kurzen Zugriffszeiten hohe Kapazitäten aufweisen, z.B. die optische Speicherplatte (DOR-Platte) (134).

Ein Teilaspekt des Mengenproblems ist das starke Anwachsen der Report-Literatur (graue Literatur). Vorzugsweise von Großforschungseinrichtungen (5,6) und anderen staatlich finanzierten Stellen produziert, stellt ihre Beschaffung und Auswertung ein Problem für das Bibliothekswesen dar.

Was kann man unter "Information" verstehen?
Aus der Literatur lassen sich verschiedene Auffassungen herausarbeiten:

- In der Informatik und Nachrichtentechnik befaßt man sich mit physikalischen und technischen Bedingungen der Informationsausbreitung und Informationsverarbeitung. Entsprechend angepaßt ist der Informationsbegriff: Informationen sind physikalische Größen, sie sind ein räumlich-zeitliches Verteilungsmuster von Materie und Energie (SHANNON, WEAVER (7)). Der apparativ-technische Aspekt steht bei dieser Betrachtungsweise im Vordergrund, man untersucht Übertragungskanäle, - geschwindigkeiten und die erreichbaren Genauigkeiten. Information wird von den Begriffen "Signal" und "Nachricht" nicht abgegrenzt.

N. WIENER (8) löst in seiner Betrachtung der Information den Begriff von konkreten physikalischen Größen wie z.B. Zeichenfolgen. Er schreibt:

"Wie der Informationsgehalt eines Systems ein Maß für den Grad der Ordnung ist, so ist die Entropie eines Systems ein Maß für den Grad der Unordnung; und das eine ist einfach das Negative des anderen."

BRILLOUIN (9) meint: "Information und physikalische Entropie sind von derselben Natur."
Hält man sich an obigen Informationsbegriff, so hat man den Vorteil der Meßbarkeit der Information. Er hat aber den Nachteil, daß er die Wirkung beim Empfänger vernachlässigt. Die Bedeutung einer Information entzieht sich den pysikalischen Meßmethoden.

Die verschiedenen Definitionen der Information leiten sich von den Aspekten ab, unter denen man die Information betrachtet.

1. Die Struktur-Definition
   Die Informationen liegen in den strukturellen Eigenschaften der materiellen Objekte. Die Bedeutung der Informationen spielt keine Rolle.

2. Die Nachrichten-Definition
   Informationen werden den Nachrichten gleichgesetzt. Da diese Definition aus Nachrichtentheorie und Kybernetik stammt, ist sie mathematisch am besten ausgebaut.

3. Die Bedeutungs-Definition
   Informationen werden als Veränderungen des Gesamtwissens aufgefaßt. Jede Veränderung, gleich ob Vergrößerung oder Verringerung, wird als Informationsprozeß aufgefaßt.

4. Die Prozeß-Definition
   Die Informationen werden als Ergebnis eines Prozesses begriffen, der aus
   - Erzeugung von Informationen,
   - der Übertragung der Informationen mittels beliebiger technischer Mittel und
   - der Wirkung der Informationen beim Empfänger besteht (22).

Den nachrichtentechnischen Aspekt der Information kann man hier vernachlässigen. Wenn technische Probleme angesprochen werden, dann nur im Zusammenhang mit der Übertragung von Signalen. Diese technischen Fragen wie z.B. die Gestaltung von Bildschirmausgaben oder die Formulierung der Suchfragen sollten benutzerfreundlich gelöst sein.

Bei der übermittelten Nachricht, die zur Information geworden ist, spielen folgende Fragen eine Rolle:
- Ist die Nachricht für die Frage relevant?
- Vermehrt sie das Vorwissen, das zur Lösung eines Problems notwendig ist?
- Trägt sie zur unmittelbaren Lösung des Problems bei?
- In welchem inhaltlichen Zusammenhang ist die Information einzuordnen?

Eine Zusammenfassung aus der Bedeutungs- und Prozeß-Definition erfüllt die gestellten Anforderungen am besten. Ein Beispiel aus der Chemie soll die Kontextabhängigkeit der Information verdeutlichen:
"Der Schmelzpunkt irgendeiner Substanzprobe beträgt 123,4 ^{o}C."

Einem Chemiker hilft dieser Satz alleine bei der Lösung eines analytischen Problems nur unwesentlich weiter, insbesondere dann, wenn er über wenig oder keine Hintergrundinformationen verfügt. Er wird nach ergänzenden Informationen suchen, die

ihm die Bewertung des Meßwertes erlauben, z.B. nach
- der Aufheizgeschwindigkeit,
- dem Meßgerät
- oder nach Angaben zur Korrektur des Meßwertes.

Hat er den Schmelzpunkt gemessen, um zwischen zwei möglichen Substanzen zu entscheiden oder um die Reinheit einer Substanz zu überprüfen und befindet sich in seinem Vorwissen eine Information über den zu erwartenden Meßwert, so kann er den Schmelzpunkt zur Entscheidung heranziehen.

Die Kontextabhängigkeit der Informationen stellt an ein Informationssystem hohe Anforderungen, sollte es doch in der Lage sein, die verschiedenen Aspekte bei der Fragestellung abzubilden. Konventionelle Speichermedien erlauben die Abfrage nach jeweils nur einem Merkmal, die gleichzeitige Suche nach mehreren Merkmalen und Kombinationen von ihnen ist eine Leistung, die durch den Einsatz von Computern möglich gemacht wird.

In der Informationswissenschaft legt man Wert auf die Bedeutung und auf die Wirkung der Information. Die Richtung, in der diese Ansätze gehen, kann man folgendermaßen andeuten: "Information ist nur, was verstanden wird." (10)
Es kommt also neben der korrekten Übertragung der Signale auch auf die intellektuelle Verarbeitung des Inhalts beim Empfänger an (11).
G. BÖHME (10) verknüpft die Struktur der materiellen Welt mit der davon getragenen Bedeutung. Nach seiner Meinung genügt es nicht, Symbole auszutauschen, sondern es muß auch gewährleistet sein, daß der Inhalt der durch die Zeichen übermittelten Nachricht verstanden wird: "Information ist stets Struktur mit Bedeutung."
Der Informationsbegriff verliert durch die Erweiterungen an Schärfe und Quantifizierbarkeit, er wird aber wirklichkeitsnäher.

Parallel zur Diskussion über "Information" verläuft die Auseinandersetzung über Gegenstand, Methodik und Inhalt der "Informationswissenschaft". Diese Auseinandersetzung (BORKO (12), SCHOBER (13), WERSIG (14), PIETSCH (15), KUNZ u. RITTEL (16), DIETSCHMANN, OCKENFELD u. PIETSCH (17), SAUNSERS (18), HECHT(19), SLAMEKA (20), GARFIELD (21) wurde lange und auch kontrovers geführt, eine pragmatische Sehweise beginnt sich durchzusetzen.

Über den Zusammenhang der Informationswissenschaft mit Nachbardisziplinen gibt das Schema ein Bild:

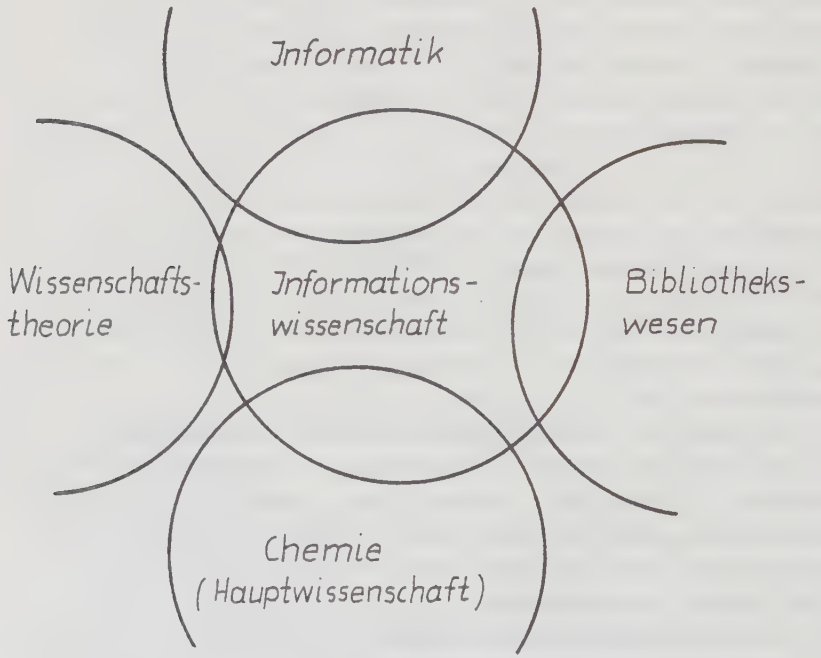


Abb. 1

1.2. Dokumentation

Wie beim Begriff "Information" taucht beim Begriff "Dokumentation" die Polysemie-Problematik auf. Unter Dokumentation versteht man das Sammeln, das Ordnen und das Nutzbarmachen von geistigen Schöpfungen jeder Art und aus allen Gebieten menschlichen Schaffens.
(Institut International de Documentation)
Nach ISO (International Standard Organization) ist Dokumentation "the collecting, organizing, storing, citing and dissemination of documents." (ISO TC 97 SC 1)

Die zweite Definition hebt zu sehr auf die Dokumente ab, auf die materiellen Träger der Informationen. Damit engt man sich unnötigerweise ein und betrachtet ungewollt nur die schriftlich niedergelegten Informationen. Es scheint daher angebracht zu sein, vom physikalischen Erscheinungsbild der Informationen zu abstrahieren und sich dem umfassenderen Dokumentationsbegriff der ersten Definition anzuschließen.

Dabei wird der Anspruch aufgegeben, ein für das gesamte menschliche Wissen geeignetes Kassifikationssystem zu schaffen (Universalklassifikation). Jeder Tätigkeitsbereich erfordert eine ihm möglichst genau angepaßte und wandelbare Struktur der Wissensordnung. Der Gefahr, die speziellen Fachbedürfnisse zu vernachlässigen, wird dadurch entgegengewirkt. Die Dokumentation wird zu einem Unterstützungsmittel der fachspezifischen Kommunikation, darüber hinaus erfüllt sie wichtige Aufgaben bei der Bereitstellung von Fachwissen für Nachbargebiete (23).

Nach Art der dokumentierten Objekte kann man folgende Dokumentationsarten unterscheiden:

- Nachweisdokumentation
  Es wird vorzugsweise Literatur zu bestimmten Autoren oder zu Sachverhalten nachgewiesen.
- Datendokumentation
  Es werden dem Benutzer in erster Linie Daten von Objekten, z.B. physikalische Daten, zur Verfügung gestellt und erst in zweiter Linie die dazugehörenden Literaturstellen. Häufig wird ihm die Analyse und die Bewertung der Daten vom Herausgeber der Datensammlung abgenommen. Die Ausführung von arithmetischen Operationen (Mittelwertbildung, Interpolation, Extrapolation und Zeitreihenbildung) und die graphische Darstellung der Ergebnisse zählen oft zum Leistungsumfang einer Datendokumentation. Es ist nicht damit getan, dem Benutzer die nackten Zahlen nachzuweisen, sondern der Systembetreiber muß auch den Kontext zu den Daten anbieten (Meßbedingungen, Labor, Substanzkennzeichnung und Fehlergrenzen) (26).

Für auf Computer implementierte Literaturinformationssysteme existieren verschiedene komfortable Abfragesprachen (GOLEM, GRIPS/DIRS, ORBIT, BRS , STAIRS etc.) die eine Abfrage der Speicher nach Kombination der Suchbegriffe (Boole'sche Logik) erlauben.
Die besonderen Anforderungen der Datendokumentation erfordern dem Zweck angepaßte Spezialentwicklungen, daher ist die Zahl der sich im Betrieb befindenden Systeme noch verhältnismäßig gering.

1.3. Kommunikation

Die fachliche Kommunikation vollzieht sich auf vielen Kanälen:

- schriftliche Kanäle
  - Bücher
  - Zeitschriften
  - Forschungsberichte (Reports)
  - Patente
  - Fortschrittsberichte
  - invisible colleges

- mündliche Kanäle
  - Gespräche
  - Konferenzen

- elektronische Kanäle
  - Datenbanken
  - Funk, Fernsehen
  - Vernetzung von Rechnern

Die fachliche Kommunikation bedient sich im wesentlichen der gleichen Kommunikationswege wie die Massenkommunikation (24, 25).
Von den Kanälen hergesehen bietet sich kein Abgrenzungskriterium an. Der Inhalt liefert das Abgrenzungskriterium. Durch Fachkommunikation werden berufs- und problemlösungsbezogene Informationen ausgetauscht, die man bei der Bearbeitung von Problemen aus Wissenschaft, Technik, Wirtschaft und Verwaltung benötigt.

Die Fachkommunikation verfügt über einige Besonderheiten:

- Es existiert eine Selbstkontrollinstanz der wissenschaftlichen Gemeinschaft in Form von Redaktionen und Herausgebergremien für Fachzeitschriften und Kongreßprogramme.
- Die Ergebnisse des Fachkommunikationsprozesses veralten nicht so schnell wie die Inhalte der Massenkommunikation und werden intensiver weiterverarbeitet und ausgewertet. Sie werden gespeichert, um für zukünftige Anfragen zur Verfügung zu stehen.

Die vorherrschenden traditionellen Wege der Kommunikation in der Wissenschaft sind u.a.:

- Zeitschriften
- Monographien
- graue Literatur
- Datensammlungen
- Patente
- Handbücher
- Vorträge
- Kongreßbesuche

Neue Fachkommunikationswege werden nur zögernd beschritten. Alle Kommunikationstechniken stehen unter dem Zwang, sich gestiegenen Anforderungen hinsichtlich Qualität und Quantität der übertragenen Informationen anpassen zu müssen.

Modelle für den Kommunikationsprozeß

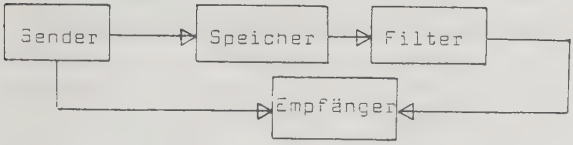


Abb. 2

Unter Einbeziehung mehrerer Speicher und der dazugehörigen Filter ergibt sich folgendes Bild.

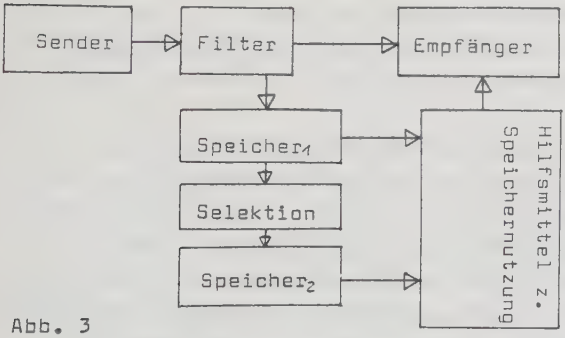


Abb. 3

Mit zunehmender Vernetzung des Kommunikationssystems entstehen zwischen Sender und Empfänger Barrieren. Ihre Überwindung ist eines der Hauptprobleme des modernen Kommunikationswesens.
Man wollte eigentlich Zugangshilfen schaffen, die es dem Benutzer ermöglichen, das gespeicherte Wissen effektiv zu nutzen. Dazu zählen Kataloge, Bibliographien und Referateorgane, die inzwischen so kompliziert und umfangreich geworden sind, daß man zu ihrer sinnvollen Nutzung wiederum eine Anleitung benötigt.
Auch informelle Kommunikation zwischen Wissenschaftlern ist inzwischen Gegenstand der Forschung, obwohl sie schwieriger als die auf den formalen Kanälen ablaufende Kommunikation zu untersuchen ist.

Bisher wurde versucht, Teilsysteme des gesamten Kommunikationsprozesses zu verbessern. Obwohl diese Vorgehensweise die Gefahr der Suboptimierung in sich birgt, ist es gerechtfertigt, da sich das Gesamtsystem in voneinander abhängige Teilsysteme zerlegen läßt.

1.4. Forschung

Wenn man unter Forschung die naturwissenschaftlich-technische Forschung versteht, dann kann man auf den von KARL R. POPPER geschaffenen Grundlagen aufbauen (27, 28).
Dieses Forschungsmodell liefert die methodologische Grundlage für die vorliegende Arbeit.

Eine aufgestellte Hypothese wird mit Hilfe experimentell gewonnener Ergebnisse überprüft. Aus den gewonnenen Daten werden Schlüsse gezogen, die die Hypothesen entweder bestätigen oder sie falsifizieren. Stimmen Ergebnis und die Vorhersage überein, so ist die Hypothese bekräftigt. Im Laufe der Zeit erfolgen immer weitere Überprüfungen, die im optimalen Fall die Arbeitshypothese immer besser bestätigen. Ein absolutes Maß an Sicherheit oder vollkommene Wahrheit kann aber grundsätzlich nicht erreicht werden, da der Ausgang der in der Zukunft liegenden Experimente zwar wahrscheinlich ebenso ausfällt wie die vorhergegangenen Überprüfungen, es aber keinen logisch einwandfreien Grund gibt, der den Schluß von der Vergangenheit auf die Zukunft rechtfertigen könnte. Dagegen wirken die Hypothese nicht bestätigende Resultate wesentlich stärker: Ein negatives Ergebnis genügt, um eine Arbeitshypothese fallen lassen zu müssen. Man wird natürlich versuchen, eine einmal aufgestellte Hypothese durch erweiternde Zusatzannahmen anzupassen, um sie in die Lage zu versetzen, weitere Phänomene

richtig zu beschreiben. Eine einmal aufgestellte Hypothese wird durch dieses Erweiterungsverfahren an Einfachheit verlieren und im Laufe der Zeit immer unübersichtlicher werden, so daß das Bedürfnis nach einer grundsätzlich neuen Theorie immer stärker wird. Je weniger die alte Theorie in der Lage ist, für die Phänomene Erklärungen zu liefern, desto dringender wird das Verlangen nach einer grundsätzlich neuen Sehweise. Diese Entwicklung führt zu einer Wellenbewegung in der Geschichte der Wissenschaften, die zuerst von KUHN (29) dargestellt wurde. Sein Ansatz erweitert die Wissenschaftstheorie um eine historische Komponente. Kuhn beschreibt die Entwicklung mit Hilfe des Begriffs des Paradigmas. Darunter versteht er einen von einer Gruppe von Wissenschaftlern zu einem bestimmten Zeitpunkt als gültig erachteten Wissenskanon, der ihnen dazu dient, die auftauchenden Fragen zu beantworten. Gleichwohl ist ihnen bewußt, daß es noch Probleme gibt, die mit dem bekannten Wissen nicht erklärt werden können. Kuhn teilt die Wissenschaft in eine "normale" Wissenschaft ein, deren Betätigungsfeld innerhalb des durch das Paradigma abgesteckten Rahmens liegt und in eine "anormale" Wissenschaft, die durch die Beschäftigung mit den Anomalien die Entwicklung neuer Theorien vorantreibt.

Durch den Vergleich der Erklärungskraft kommt man dazu, das alte Paradigma aufzugeben und das neue Paradigma zu bezutzen. Diesen Vorgang bezeichnet Kuhn als eine wissenschaftliche Revolution.

Die Ansätze von Popper und Kuhn beschreiben die kollektive Entwicklung der Wissenschaft. Aussagen über die bei einem einzelnen Forscher ablaufenden Denkvorgänge liefern sie nicht. Sie machen insbesondere keine Aussage über das Entstehen von neuen Ideen, sie beschränken sich auf die Entwicklung von Mechanismen, die dazu dienen können, Ideen auf ihren

Wahrheitsgehalt zu überprüfen. Der schöpferische Vorgang beim Wissenschaftler bleibt bei ihnen unberücksichtigt. Nahm man früher einfach an, daß dem schöpferischen Akt eine wichtige Rolle in der Wissenschaft zukommt, so neigt man heute mehr der Ansicht zu, daß der Forschungsprozeß aus einer Vielzahl von genau beschreibbaren einzelnen Problemlösungsschritten besteht. Nachdem man z.B. ein Experiment durchgeführt hat, werden die Daten interpretiert. Zu diesem Zweck benötigt man eine Auskunft darüber, ob die Daten in gleicher oder ähnlicher Form schon vorliegen oder ob sie mit den Vorhersagen aus dem verwendeten Modell übereinstimmen. Jede der Prüfungen erfolgt schrittweise nach dem von KUNZ und RITTEL (30) angegebenen Schema. Mit Hilfe dieses Instrumentariums kann man zwar den eigentlichen schöpferischen Akt nicht näher untersuchen, man kann aber den Forschungsprozeß in beschreibbare Einzelschritte zerlegen, d.h. in eine Folge von Such- und Problemlösungsprozessen, um daraus Schlüsse zur Unterstützung des Problemlösungsverhaltens der Chemiker zu ziehen.

Die Entwicklung der Schachcomputer soll diese Auffassung verdeutlichen:

- Es gibt noch keine Maschine, die in der Lage wäre, eine Spielstellung so zu analysieren, wie sie ein menschlicher Spieler bewertet. Die Überlegenheit des Computers beruht nur darauf, daß er in der Lage ist, die möglichen Züge schneller als ein Mensch vorauszuberechnen. Die grundsätzliche Überlegenheit des Menschen im Analysieren von Stellungen wird immer mehr durch die schnelle, aber sture Rechenfähigkeit der Maschine kompensiert (31, 32).

Problemlösende Programme in der Chemie haben größere Schwierigkeiten zu überwinden: Es gibt mehr Regeln in der Chemie und zudem sind sie nicht so exakt formulierbar wie beim Schach. Wenn man also in den Forschungsprozeß fördernd

eingreifen will, so muß man folgende Bedingungen beachten:
- Man kann nur Teilaufgaben lösen und
- logische Modelle zur Simulation von Teilprozessen sollten verwendet werden.

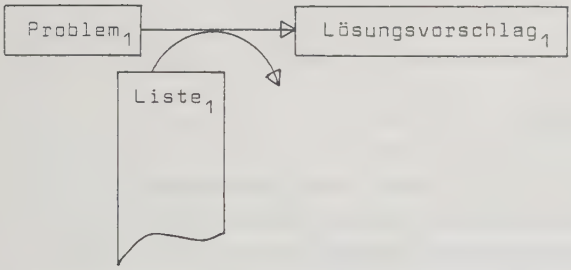


Abb. 4

Auf ein gestelltes Problem gibt eine Sammlung von relevanten Daten Auskünft. Dabei ist es gleichgültig, auf welche Art und Weise diese Liste erstellt wird: ob durch direkte Beobachtung, durch Experimente oder durch einen Vergleich mit bereits vorhandenen Daten.

Folgende Methoden zur Informationsbeschaffung werden berücksichtigt:
- Suche nach den benötigten Daten, Fakten oder Sachverhalten in der Literatur,
- Berechnung der Werte mit Hilfe von Modellvorstellungen und
- Messung der Daten und Vergleich der Ergebnisse mit den Vorhersagen von Modellen.

Die zuletzt erwähnte Methode wird in diesem Zusammenhang nicht weiter diskutiert.

Die zweite Methode ist an das Vorliegen von zuverlässigen Meßdaten und Berechnungsmethoden gebunden, so z.B. bei Modellrechnungen in der Quantenchemie oder bei Berechnungen in der chemischen Thermodynamik, wo die makroskopischen Eigenschaften von Reinstoffen und Gemischen zum Zwecke der technischen Verarbeitung bekannt sein müssen (33 - 38).

1.5. Informationssysteme

In diesem Zusammenhang werden unter "Informationssysteme" alle Einrichtungen verstanden, die dazu dienen, ein bestehendes Informationsbedürfnis zu befriedigen, unabhängig davon, ob das Angebot an Information gedruckt oder maschinenlesbar - bewertet oder unbewertet - aus Hinweisen auf Quellen oder in unmittelbarer Sachverhaltsinformation besteht.

1.5.1. Die Handbücher

Für die organische und anorganische Chemie bestehen zwei traditionsreiche Handbücher:
- der "Beilstein" und
- der "Gmelin",

benannt nach den Gründern dieser Werke (39 - 41). Hauptaufgabe beider Handbücher ist die Sammlung, die Bewertung und Zusammenführung von Daten und Fakten zu Verbindungen. Alle Daten werden an einer Stelle in der logischen Systematik publiziert. Die Erstellung bedingt Verzugszeiten bis zu etwa fünf Jahren. Eine gewisse Zeit muß man verstreichen lassen, um der wissenschaftlichen Gemeinschaft Gelegenheit zu geben, auf publizierte Ergebnisse zu reagieren. Wichtige Ergebnisse und Richtigstellungen werden früher veröffentlicht (42, 43).

Es laufen Versuche, den Inhalt von Handbüchern maschinenlesbar zu erfassen und ein Retrievalsystem zu entwickeln, das dem Benutzer die manuelle Suche erleichtert und die maschinelle Suche ermöglicht (44).
Als Ergebnis der Suche im Handbuch erhält man eine Liste der geprüften Daten und die Literaturstellenhinweise auf die Quellen der Daten. Die Auskünfte sind oft so ausführlich, daß man die Originalliteratur nicht mehr einsehen muß.
Durch diesen großen Vorteil werden die kritisch erstellten Handbücher auch in Zukunft ihren Platz in der Informationslandschaft behaupten.
Es wird angenommen, daß der unkritische Nachweis aller relevanten Literaturzitate den Benutzerbedürfnissen immer weniger entsprechen wird. Der Wissenschaftler ist in der Lage, die Originalliteratur zu bewerten, der Techniker ist auf gesicherte Daten angewiesen.
Um einen Eindruck vom Umfang der Bewertungsarbeit zu vermitteln, sei eine Zahl aus dem Beilstein-Institut zitiert: etwa 90% der Kosten für einen Band des Beilstein-Handbuches entstehen bis zur Herstellung des druckfertigen Manuskripts. Satz, Druck, Einband, Vertrieb, Werbung und Verlagsgewinn machen 10% der Gestehungskosten aus (54).

1.5.2. Literaturinformationssysteme

Chemical Abstracts

Das umfassendste Referateorgan der Chemie, Chemical Abstracts (CA) bemüht sich um die möglichst vollständige Erfassung all der Literatur, die für die Chemie und ihre Nachbargebiete einschlägig ist.

Zu diesem Zweck werden etwa 14.000 Zeitschriften überwacht und ca. 500.000 Referate pro Jahr veröffentlicht (Werte 1980). Der Umfang der Veröffentlichungen (Referate, Sachregister, Namensregister, Summenformelregister und Patentkonkordanz) ist inzwischen so umfangreich geworden, daß besondere Einführungen und Register für die Register erscheinen (Index Guide und Parent Compound Handbook).
Ein Handbuch leitet zum Gebrauch der gedruckten Dienste an (45). Eine erfolgreiche manuelle (und maschinelle) Suche in den Registern setzt eine gute Kenntnis der Indexierungspraxis von CA voraus, um nicht durch unpassende oder unpräzise Formulierung der Frage einen Informationsverlust zu erleiden (46).

Schematischer Aufbau der CA

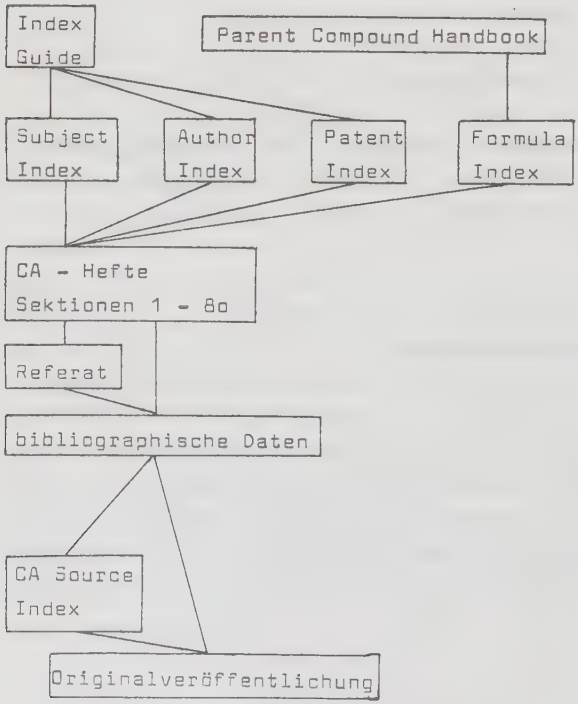


Abb. 5

Science Citation Index

Ein einzigartiges Suchsystem bietet der Science Citation Index an. Mit Hilfe der Zitatennachweise ist man in der Lage, von einer bekannten Literaturstelle ausgehend bei der Suche zeitlich vorwärtsschreitend zu arbeiten. Das ist bei keinem anderen Informationssystem möglich. Der SCI liefert mit seinen Daten das Grundmaterial für die Zitatenanalysen (47, 48).
Es werden ca. 2.000 naturwissenschaftliche Zeitschriften ausgewertet, die Daten werden zu Registern verarbeitet. Als Suchmerkmal dient der Name des Autors, man findet die Namen der Kollegen, die den Autor mit seiner Arbeit zitiert haben. Ein permutierter Titelindex gibt einen Anhaltspunkt ob der Inhalt der gefundenen Arbeit für die Fragestellung relevant sein könnte.
Die Suche ist mit einigen Fehlermöglichkeiten behaftet:

- Populäre Autoren werden häufiger gefunden.
- Gefälligkeits- und Selbstzitate verfälschen das Bild.
- Autoren, die in weniger häufig gesprochenen Sprachen schreiben, werden unverhältnismäßig wenig zitiert.

Mit dem Material, das im SCI vorliegt, kann man die folgenden Fragen beantworten:

- Welcher wissenschaftliche Autor wird in einem bestimmten Zeitraum am häufigsten zitiert?
- Gibt es auffallend häufige Zitierungen, die auf einen sachlichen Zusammenhang schließen lassen?
- Kann man solche Häufungen als Beleg für die Herausbildung einer wissenschaftlichen Schule interpretieren?

1.5.3. Maschinenlesbare Dienste

Da in zunehmendem Maße vom maschinenlesbaren Magnetband gedruckt wird, liegen viele Publikationen in zwei Versionen vor:

- in einer gedruckten Fassung, die auf den konventionellen Wegen verbreitet wird und
- in einer maschinenlesbaren Version, die ursprünglich zur Herstellung der Druckvorlage diente, jetzt aber zunehmend direkt genutzt wird, um die Informationsversorgung zu beschleunigen.

Die Speicherung von Volltexten einer wissenschaftlichen Publikation stößt noch auf große Schwierigkeiten. Die dazu notwendigen Massenspeicher sind noch zu teuer, die Retrievalverfahren sind noch nicht schnell genug, das Urheberrechtsproblem wirft noch Fragen auf, und nicht zuletzt ist die Benutzerschaft noch skeptisch.
Die linguistischen und semantischen Probleme bei der Suche in Volltextspeichern sind noch nicht befriedigend gelöst. Daher verwundert es nicht, daß zuerst die Referateorgane anfingen, ihre Dienste auch maschinenlesbar anzubieten. Eignen sich doch die Register mit ihren Datenstrukturen besonders für eine maschinelle Abfrage.
Während man bei der manuellen Suche in den Registern immer nur nach einem Merkmal suchen kann, erlauben es die Retrievalprogramme, die Abfrage nach mehreren Merkmalen durchzuführen. Die maschinelle Recherche ist spezifischer und vollständiger, als es eine manuelle Suche je sein könnte (49).

Die wichtigste maschinenlesbare Datenbasis für Chemiker stellt - analog zu der Situation bei den gedruckten Diensten - Chemical Abstracts Service (CAS) zur Verfügung (50).

Die Literaturdaten vertreibt CAS allerdings nicht selbst, sondern überläßt die Vermarktung mehreren kommerziellen Datenbankbetreibern, die die Daten im Rahmen ihres Gesamtprogramms anbieten.
Die maschinelle Verarbeitung der rd. 5.000.000 Strukturformeln stellte lange Zeit ein großes Problem dar, da die Kapazität der Rechner nicht ausreichend war, um so große Informationsmengen zu verarbeiten (51).
CAS bietet jetzt auch die Recherche nach Strukturformeln an und bedient sich folgenden Kunstgriffs:
Anstelle eines großen Rechners werden mehrere leistungsfähige Mikrorechner eingesetzt, die immer einen Teil der gespeicherten Strukturformeln durchsuchen (CAS-online).

Schematischer Aufbau der Strukturrecherche bei CAS

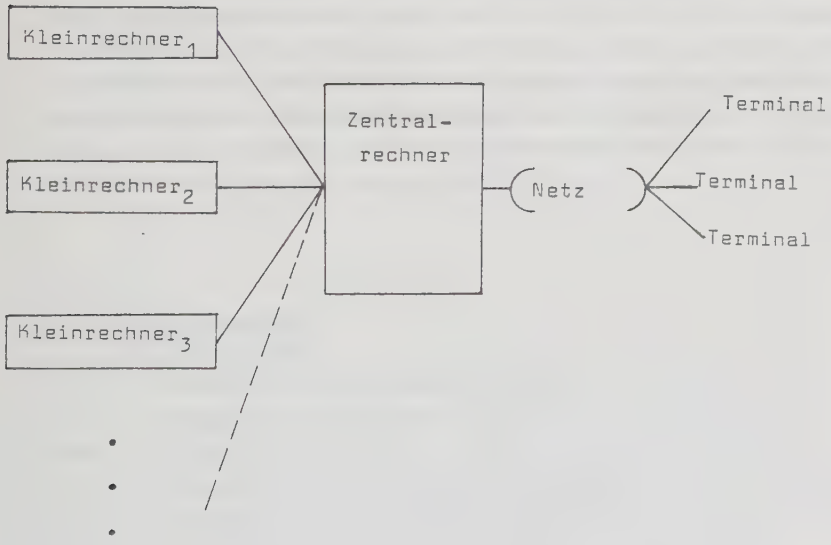


Abb. 6

Wächst die Menge der zu verarbeitenden Strukturformeln über die Kapazität der gerade benutzten Rechner hinaus an, so wird die Hardware-Ausstattung um einen weiteren Rechner ergänzt.
Eine Anfrage wird im Stapelbetrieb verarbeitet, die gefundenen Ergebnisse werden dem Benutzer sofort mitgeteilt, so daß die Antwortzeiten kurz sind.
Als Rechercheergebnis werden CA-Registry-Nummern geliefert, dazu die zehn neusten Literaturhinweise, die Summenformeln und die Formelbilder.

Bei der Beschreibung der Leistungsfähigkeit der maschinellen Recherche in Literaturdatenbanken sollte nicht vergessen werden, daß das Rechercheergebnis aus Hinweisen auf Veröffentlichungen besteht. Dem Benutzer ist es überlassen, sich die Originaldokumente zu beschaffen und auszuwerten. Literaturinformationssysteme streben Vollständigkeit an und bewerten die Literatur im Gegensatz zu den Handbüchern nicht. Es zeigt sich, daß der Benutzer in immer stärkerem Maße neben oder anstelle von Literaturnachweisen direkte Informationen verlangt. Dies hat zur verstärkten Entwicklung von Fakten- und Dateninformationssystemen geführt.

1.5.4. Fakten- und Dateninformationssysteme

Dateninformationssysteme weisen dem Benutzer die Daten direkt nach, ohne ihn erst auf die Literatur zu verweisen (68).

Obwohl für die Dienstleistungen von Fakten- und Dateninformationssystemen große potentielle Nachfrage besteht, existieren erst wenige, spezielle Systeme.
Folgende Gründe sind dafür maßgebend:

- Die Beschaffung, Auswahl und Einspeicherung der Daten ist ungleich aufwendiger als die Eingabe bei Literaturinformationssystemen.
  Ohne eine Bewertung der Fakten sollte man ein System nicht aufbauen.
- Es wurden erst wenige Retrievalsysteme entwickelt; die beim Retrieval anfallenden Rechenzeiten sind hoch und führen zu Diskrepanzen zwischen dem Ergebnis und den dafür aufzuwendenden Mitteln.
- Mit einer einzelnen Zahl ist dem Benutzer selten gedient, er verlangt oft nach Tabellen oder graphischen Darstellungen, die als Entscheidungsgrundlagen dienen können. Dieser Komfort verlangt ein Software-Modul und ein Terminal für graphische Datenverarbeitung (52, 53).

1.5.5. Referral-Systeme

Die Vielzahl bestehender Informationssysteme und deren wachsende Autonomie erfordern neue Konzepte für den Verbund der Systeme. Die Abfrage von mehreren Datenbanken zu einer Fragestellung (file-switching) ist aber nur dann möglich, wenn Umfang, Struktur, inhaltliche und formale Merkmale, Indexierungspraxis und Zugangsmöglichkeiten der einzelnen Informationssysteme bekannt sind.

Zu diesem Zweck existieren Informationsvermittlungsstellen, die möglichst viele der relevanten Quellen aus fachlicher und organisatorischer Sicht kennen und auswählen können, so daß dem Anfrager eine angemessene Antwort gegeben werden kann.
Es werden alle Informationssysteme, unabhängig, ob sie Literaturhinweise, Daten, Fakten, Expertennamen oder Produktinformationen liefern, befragt.

1.6. Spezielle Chemie-Informationssysteme

Da die Chemical Abstracts das Gesamtgebiet der Chemie erschließen, kommt den übrigen Referateorganen in der Chemie die Rolle der Ergänzungsmedien zu.
Kleinere Informationsdienste haben durch ihr benutzerspezifisches Angebot einen wichtigen Platz in der Informationsversorgung der Chemiker.
Die wichtigsten deutschsprachigen Publikationen sind:

- Chemischer Informationsdienst (ChemInform)
  Der ChemInform beschreibt die Reaktionen aus der niedermolekularen organischen und der anorganischen Chemie durch Strukturformeldiagramme.
- Das Konzentrat
  Dieser Dienst wertet Publikationen aus dem Gebiet der chemischen Technik und der Biotechnologie aus.
- Literaturschnelldienst Kunststoffe Kautschuk Fasern
  Die makromolekulare Chemie wird von diesem Dienst abgedeckt.
- Chemisches Zentralblatt
  Für Literaturrecherchen vor 1969 (Einstellungsdatum des Referateorgans) sind die Bände mit ihren informativen und ausführlichen Referaten gut geeignet.

Strukturformeldarstellungen

Chemische Sachverhalte kann man verbal oder mit Hilfe der Strukturformeln beschreiben. Die Diagramme stellen zweidimensionale Modelle für die dreidimensionale Realität der Moleküle dar. Um die Strukturformeln für Computer verarbeitbar zu machen, werden folgende Wege beschritten:

Linearnotationen

Die bekannteste Linearnotation ist die Wiswesser Line-Notation (WLN). Sie verwendet Buchstaben und Sonderzeichen einer Schreibmaschinentastatur zur Codierung der Strukturen. Das Verfahren beginnt an dem Teil der Verbindung, dessen Symbol in der Schlüsseltabelle an der letzten Stelle steht. Die weiteren Elemente werden in der Reihenfolge zugeordnet, in der sie im Molekül miteinander verbunden sind.

Fragmentcodes

Fragmentcodes für den Bereich der Strukturverarbeitung stellen ein gut ausgebautes Hilfsmittel für die Strukturrecherche dar, sind aber mit dem Nachteil behaftet, einen hohen Aufwand bei der Einspeicherung und bei der Recherche zu erfordern, da man eine neue formale Sprache mit Vokabeln (= Verschlüsselungselemente) und einer Grammatik (= Verknüpfungsregeln für die Vokabeln) lernen muß. Wenn man die Sprache beherrscht, stellen die Fragmentcodes eine einfache Möglichkeit dar, nach Verbindungen und insbesondere nach Verbindungsklassen zu fragen. Ein grundsätzlicher Nachteil haftet allen Fragmentcodes an: sie können nur die Strukturen verschlüsseln, die bei ihrer Konzeption bekannt waren, um berücksichtigt zu werden. Im Lauf der Zeit entsteht ein immer größeres Bedürfnis, die Codes zu erweitern und damit zu komplizieren. Im konsequent durchdachten Extremfall

kommt man dann zur Verschlüsselung für jedes einzelne Atom und damit zu einer topologischen Verschlüsselung (55,56). Mit den Fragmentcodes kann man nach Verbindungsklassen recherchieren. Topologische Codes ermöglichen die Suche nach jeder einzelnen Verbindung. Ihren wichtigen Platz werden die Fragmentcodes in der Patentdokumentation behalten, da man durch Fragmentcodes allgemeine Formulierungen verschlüsseln kann. Die Doppelnatur der Patente als juristische und wissenschaftliche Dokumente bringt es mit sich, daß Ansprüche auf Patentschutz sehr allgemein formuliert werden. Durch die Möglichkeit der Klassenbildung lassen sich Patentrecherchen durch Fragmentcodes durchführen. (GREMAS-Code)

Topologische Darstellungen

Das Strukturformelrecherchesystem DARC (Description, Acquisition, Retrieval, Correlation) mit graphischer Eingabe basiert auf der Topologie von Chemical Abstracts. Die Verarbeitung der Anfrage beruht auf dem Konzept der Strukturcluster.
Um einen Focus herum ist eine Umgebung definiert, die stufenweise erweitert wird und sich auf den Focus zentriert. Dieses Beschreibungselement nennt DUBOIS (58) ELCO (Ordered Concentric Limited Enviroment), es kann in einem iterativen Prozeß zur Beschreibung des vollständigen Moleküls ausgebaut werden:

$$(FO)^{n+1} = (FO)^{n} + E^{n}$$

$(FO)^{n+1}$: neuer Focus
$(FO)^{n}$: alter Focus
E^{n} : Enviroment um den Focus $(FO)^{n}$

Die Entwicklung der beiden Retrievalsysteme von Chemical Abstracts und von CNRS (DARC) befindet sich noch stark im Fluß, so daß es noch nicht möglich ist, bewertende Aussagen über die Leistungsfähigkeit der Systeme zu machen (57, 58).
Einen Nachteil haben allerdings beide Recherchemethoden: sie sind teuer, wenn man nur die unmittelbar anfallenden Kosten in Betracht zieht. Berücksichtigt man aber, daß man seine Recherchefrage wesentlich treffsicherer formulieren kann, wenn man Strukturformeln oder Strukturformelfragmente als Beschreibungsmerkmale verwenden kann, so wird deutlich, daß sich der Aufwand lohnt.

2. Syntheseplanungssysteme

Diese besonders fortschrittlichen Systeme können dem Forschungsgebiet künstliche Intelligenz (KI) zugeordnet werden, dessen theoretische und methodologische Grundlagen aus folgenden Bereichen stammen:

- Logik und Mathematik
- Informatik und Signalverarbeitung
- Graphentheorie (62)
- Psychologie und
- Semantik

KI befaßt sich mit den Methoden zum Erwerb, zur Darstellung und zur Verwendung von Wissenselementen mit Unterstützung durch EDV. Ansatzpunkt der Bemühungen zur Simulation der Denkprozesse ist die kognitive Psychologie, die untersucht wie Menschen sich verhalten, wenn sie Probleme wahrnehmen und bearbeiten (59).

KI-Methoden finden auf folgenden Gebieten Anwendung:

- Spieltheorie (Schachprogramme)
- Mathematik (Programme, die mathematische Sätze ableiten und beweisen)
- Linguistik (Programme zur automatischen Übersetzung und zum automatischen Indexieren)
- Bildverarbeitung und Mustererkennung
- graphische Datenverarbeitung und Computerkunst
- Simulationsmodelle
- Modelle zur Repräsentation des Wissens
- Systeme in der Psychologie (Programm ELIZA) (59)

Ein früher Versuch, die Methoden der KI auf chemische Fragestellungen anzuwenden, ist ein Programm zur automatischen Interpretation von Massenspektren.
Das Programm "DENDRAL" von I. LEDERBERG (60, 61) ist modular aufgebaut:

1. Planungsmodul: Aus den vorhandenen Angaben werden notwendige Einschränkungen formuliert. Dieser Schritt ist notwendig, um die Menge der in den folgenden Schritten zu untersuchenden Alternativen möglichst gering zu halten.
2. Erzeugungsmodul (CONGEN) Hier werden systematisch die vollständigen Listen der Strukturen erzeugt, die aufgrund der Einschränkungen noch erlaubt sind.
3. Programmteil: Hier wird für jede Struktur ein Massenspektrum berechnet. Das vorhergesagte Spektrum wird mit dem Eingabespektrum verglichen. Strukturvorschläge, die zu übereinstimmenden Spektren führen, werden als richtige Antworten ausgegeben.

Mit DENDRAL und der Weiterentwicklung HEURISTIC DENDRAL können Massenspektren von Ethern, Ketonen, Aminen und Alkoholen zugeordnet werden.
Wenn systematisches Absuchen und Erzeugen von prinzipiell möglichen Problemlösungen gefragt sind, so ist DENDRAL dem Fachmann überlegen, da es vollständiger und systematischer arbeitet. Seine Leistungsfähigkeit erstreckt sich nur auf den vorher programmierten Bereich.
Auch auf dem Gebiet der Arzneimittelplanung finden Methoden der KI Eingang (Computer aidet drug design = CAD). Um eine Beziehung zwischen der Struktur und dem Wirkungsspektrum einer Substanz herzustellen, wendet man Cluster-Verfahren und Pattern-recognition-Prozesse (Mustererkennungs-Prozesse) an. Die Suche nach aussichtsreichen Synthesevorschlägen wird beschleunigt (137).
Die Möglichkeiten der Syntheseplanungssysteme reizen zum Vergleich mit der Leistungsfähigkeit der menschlichen Problemlöser (63 - 65).
Neben der direkten Funktion der Unterstützung des Chemikers bei seiner Arbeit durch Verlagerung einiger Funktionen des Gedächtnisses, durch Anregung der Phantasie und durch weitgehende Abarbeitung eines Syntheseproblems trägt die Beschäftigung mit Syntheseplanungssystemen zur sorgfältigen Problemanalyse und zur Beschäftigung mit den Regeln und Gesetzmäßigkeiten der Syntheseplanung bei. Die dabei aufgefundenen Denkprinzipien lassen sich bei der herkömmlichen Planung von Synthesen anwenden.
Viele der von Corey (66) publizierten Regeln waren den Chemikern vorher schon bekannt, man hatte sie sozusagen im Gefühl. Um aber einen Rechner mit diesen Regeln programmieren zu können, war es notwendig, die Regeln und ihren Gültigkeitsbereich zu formulieren.

Für Syntheseplanungssysteme lassen sich zwei verschiedene Planungsansätze unterscheiden:

- aufbauend auf einer Liste von Stoffen und von Reaktionen wird ein Spezialgebiet der Chemie abgedeckt. Diese induktive Arbeitsweise hat den Vorteil, recht schnell zu ersten Ergebnissen zu führen, erfordert aber viel Aufwand zur Systempflege und Wartung.
- Das System von UGI und seinen Mitarbeitern (67) versucht, einen deduktiven Ansatz zu verfolgen. Die Vorschläge werden durch ein mathematisches Modell der Chemie generiert.

Jedes Syntheseplanungssystem fungiert gleichzeitig als Produzent und Verbraucher von Wissen:
Verbraucher von Wissen ist es, wenn es

- auf die in der Reaktionsbibliothek gespeicherten Reaktionen zurückgreift, um Umsetzungen durchzuführen,
- auf die Stoffbibliothek zurückgreift, um eine Syntheseplanung abzuschließen oder um passende Reaktionspartner zu finden

Produzent von Wissen ist es, wenn es

- die Grenzen der Gültigkeit der Regeln an speziellen Fällen aufzeigt und so Anlaß gibt, die Regeln zu erweitern,
- neue, formal zwar mögliche, aber chemisch zumindest ungewöhnliche Synthesewege aufzeigt, z.B.:

Abb. 7

- neue Beziehungen zwischen Eigenschaften und Strukturen aufzeigt.

J. H. WINTER gibt eine knappe Übersicht über die Syntheseplanungssysteme von Corey, Wipke, Gelernter, Bersohn, Ugi, Moreau und Weise (135).

2.1. Auf einer Stoff- und Reaktionsbibliothek beruhende Systeme

Als Vorläufer dieser Systeme kann DENDRAL gelten, das sich mit der Vorhersage chemischer Reaktionen unter den Bedingungen des Massenspektrometers befaßt.
Vereinfacht dargestellt, geht der Syntheseplaner von ähnlichen Strukturelementen im Ausgangs- und Zielmolekül aus und versucht, sie durch geeignete Reaktionen miteinander zu verknüpfen. Dabei bewegt er sich auf einem Reaktionspfad sowohl worwärts als auch rückwärts, d.h. es findet eine wechselseitige Anregung statt zwischen den Erfordernissen der Strukturumwandlung und den bekannten Reaktionen, die zum Ergebnis haben, daß neue Reaktionen entwickelt werden können. Das maschinengerechte Verfahren ist das der Retroanalyse. Man erreicht auf diese Weise eine schrittweise Vereinfachung des Problems, solange, bis man zu einer Verbindung gelangt, die in der "Stoffbibliothek" verzeichnet ist. Es wird das Zielmolekül vorgegeben. Die Analyse des Zielmoleküls ergibt nun eine bestimmte Anzahl von Vorläufermolekülen der ersten Generation. Jedes vom Programm erzeugte Vorläufermolekül wird dann in der gleichen Weise zerlegt wie der Ausgangsstoff. Dieser Prozess wird solange fortgeführt, bis eines der erzeugten Vorgängermoleküle in der Liste der Startmoleküle gefunden wird.

Es ist nun die Aufgabe des experimentell arbeitenden Chemikers, die maschinell erzeugten Synthesevorschläge auf ihre Durchführbarkeit in der Praxis zu beurteilen.

Schema des Programmablaufs:

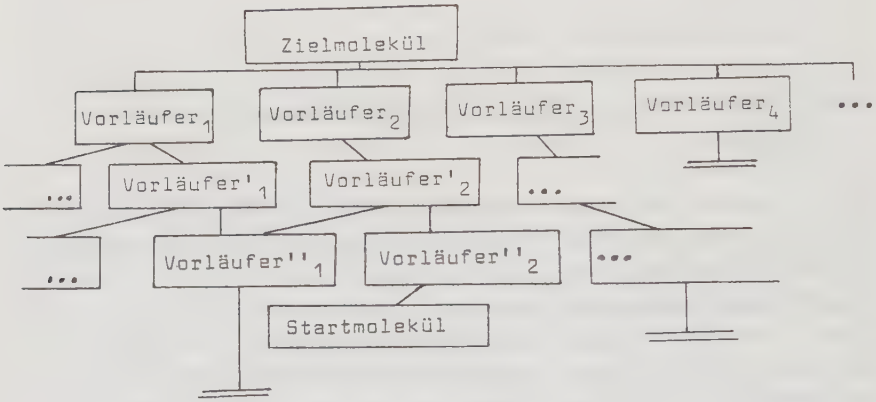


Abb. 8

Die Kompliziertheit und der Umfang des Synthesebaums sind nur durch die Kapazität der zur Verfügung stehenden Rechenanlage und durch die Rechenzeiten begrenzt. Die Startmoleküle haben auf die Richtungen im Synthesebaum keinen Einfluß, sie dienen als Abbruchbedingungen für die Suche. Alle aufgefundenen Synthesewege sind prinzipiell gleichwertig. Als Bewertungsmaßstäbe können die Energiebilanzen der benutzten Reaktionen dienen, die in einer Untereinheit des Programms berechnet werden oder Güteziffern für jede Reaktion, die neben der Energiebilanz auch kinetische oder sterische Faktoren berücksichtigen.
Das Verfahren der retrosynthetischen Analyse erscheint auf den ersten Blick ungewöhnlich. Da aber für jeden Teilschritt die mechanistische Behandlung von Hin- und Rückreaktion gleich ist, - es gilt das Prinzip der mikroskopischen Reversibilität - ist diese Vorgehensweise gerechtfertigt.

Als Ziele in der Synthesestrategie kommen in Betracht:
- die Einführung von funktionellen Gruppen,
- die Umwandlung von funktionellen Gruppen, um dadurch Modifikationen in der Reaktionsfähigkeit zu erhalten, und
- Umlagerungen

Um das Programm möglichst komfortabel zu halten, hat der Benutzer während der Durchführung der Syntheseplanung Eingriffsmöglichkeiten in den Ablauf.
Durch den modularen Aufbau ist das Programm leicht zu ändern, ohne den Aufwand zu groß werden zu lassen.
Das Graphik-Modul besteht aus einer Steuereinheit für
- Bildschirme
- Rand-Tablett mit Lichtgriffel und
- Drucker.

Dadurch ist der Chemiker in der Lage, direkt, d.h. ohne Codes mit der Maschine zu kommunizieren. Das Graphik-Modul nimmt ihm die Übersetzungsarbeit in eine Kunstsprache ab. Im Wahrnehmungs-Modul werden die vom Graphik-Modul gelieferten Informationen verarbeitet. Erkannt werden Ringe und Ringgröße, Ketten und funktionelle Gruppen, weiter noch das Molekül vereinfachende Symmetrieelemente.
Die vom Graphik-Modul aufbereiteten und vom Wahrnehmungs-Modul vorbereiteten Informationen werden im Modul für Strategie umgesetzt. Die Synthesestrategien basieren auf heuristischen Regeln der organischen Chemie für
- funktionelle Gruppen
- das C-Gerüst und
- die Stereochemie

Beispiele für Heuristiken:
- Gibt es ähnliche Fragestellungen?
- Gibt es im Molekül Symmetriezentren, die das Molekül zerlegen lassen?

- Ist das Molekül aus mehreren einfach Fragmenten zusammengesetzt?
- Welche sind die strategisch wichtigen Bindungen im Molekül? Sie sollten so spät wie möglich geknüpft werden. Solche Bindungen sind:
  - Ringbindungen in 5 - 7 gliedrigen Ringen
  - exo-Bindungen bei größeren Ringen
  - Bindungen, die nicht zwei Ringen gemeinsam sind, die zusammen mehr als 7 C-Atome haben
  - endo-ständige Bindungen in Bezug auf einen Ring, der mit möglichst vielen anderen Ringen verknüpft ist
  - Bindungen, nach deren Spaltung keine asymmetrischen C-Atome entstehen und
  - Bindungen, deren Spaltung die Anzahl der Ringe vermindert.
- Labile funktionelle Gruppen sind möglichst spät einzuführen,
- bei der Wahl des Vorläufers einer Verbindung ist so vorzugehen, daß möglichst viel Funktionalität gebildet wird,
- Reaktionen, die möglichst wenig Veränderungen bringen, sind bevorzugt anzuwenden,
- wenn man eine strategische Bindung knüpfen will und es entsteht ein Hindernis, so ist zuvor dieses Hindernis zu entfernen.

Wertvolle Reaktionen in diesem Sinne sind:
- Diels-Alder-Reaktion
- Robinson-Anellierung
- Birch-Reduktion und
- säurekatalysierte Olefincyclisierungen

Weitere heuristische Regeln befassen sich mit Eigenschaften von π-Elektronengruppen. Bei den Heuristiken ist auffallend, daß sie sich im wesentlichen mit dem C-Gerüst einer Verbindung beschäftigen und auch Einflüsse berücksichtigt werden, die von funktionellen Gruppen auf das Kohlenstoffgerüst ausgeübt werden. Die Chemie der heterocyclischen Verbindungen wird vorerst nicht verarbeitet.

Die Verarbeitung der Reaktionen erfolgt im Manipulations-Modul.
Beispiel: Die Retro-Diels-Alder-Reaktion

Abb. 9

Zuerst fragt das Programm nach der Anzahl der Ringglieder. Ist sie gleich sechs, wird nach der Doppelbindung gesucht. (Fehlt die Doppelbindung, so wird ein Unterprogramm aktiviert, das zum Ziel hat, die Doppelbindung einzuführen.) Liegt eine Cyclohexansubstruktur vor, so kann das Programm die Transformation durchführen.
Nach diesem Schema werden folgende Reaktionen behandelt:

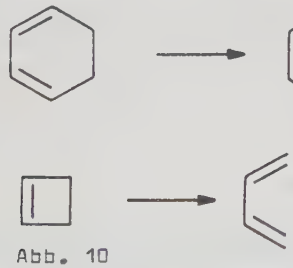


Abb. 10

Die Bewertung der vom Programm gefundenen Lösungen erfolgt durch den Chemiker und dann durch das Programm.
Maschinelle Vortests prüfen, ob durch die Manipulationen nicht unsinnige Ladungsverteilungen entstanden sind oder unmögliche Topologien, weiter wird geprüft, ob durch die Anwendung der Transformationen nicht wieder die Ausgangssubstanz erhalten wurde. Die mehr formalen Prüfungen werden durch die Berechnung von Bewertungskriterien für die erzeugten Strukturen, Ringsysteme und die daran gebundenen Gruppen ergänzt.

Das Bewertungs-Modul bewältigt drei Aufgaben:
- es verwirft offensichtlich ungenügende Strukturvorschläge,
- es mißt den Grad der Annäherung an das Ziel und
- es liefert Hinweise für die weitere Erkundung des Synthesebaums.

Es ist offensichtlich, daß die Qualität eines Syntheseplanungssystems wesentlich durch die Qualität und die Quantität der verwendeten Stoff- und Reaktionsbibliothek bestimmt ist. Damit hat man zwar die Möglichkeit, durch Anpassung an aktuelle Probleme den wechselnden Aufgabenstellungen gerecht zu werden, man ist aber gezwungen, beträchtlichen Aufwand in die Systementwicklung und in die Systempflege zu investieren. Durch Kombination mit Berechnungsmodulen für chemische, physikalische und pharmazeutische Eigenschaften kann man sehr einfach Ergebnisse erhalten, die auch in der industriellen Forschungspraxis relevant sind (72 - 74).

2.2. "Voraussetzungslose" Systeme

Mit Hilfe eines mathematischen Modells der Chemie wird versucht, chemische Probleme ohne Rückgriff auf schon vorhandenes experimentelles Einzelwissen zu lösen.

Das Konzept basiert aus folgenden Voraussetzungen: (67)
- Die eingesetzte Materie bleibt bei einer chemischen Reaktion erhalten, sie wird durch die Valenzelektronen repräsentiert. Man beschränkt sich auf die Chemie, die der Oktett-Regel gehorcht.
- Der klassische Isomeriebegriff wird erweitert. Bezog sich diese Definition auf die verschiedenen Möglichkeiten der Atomanordnungen in einem Molekül, so betrachtet man nun Ensembles von Molekülen und definiert sie als eine Gruppe

von Molekülen, die durch eine Reaktion ineinander übergeführt werden.
Isomer im Sinne von Ugi sind z.B. folgende Ensebles von Molekülen:

oder auch:

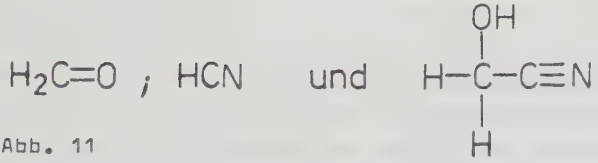


Abb. 11

Zu einem gegebenen Atomsortiment lassen sich nun viele mögliche Atomanordnungen bilden, diese Anordnungen bilden die Familie der isomeren Ensembles von Molekülen (FIEM).
Die Atomanordnungen werden durch die Bindungselektronenmatrizen beschrieben.

Beispiel: Bindungselektronenmatrix des Formaldehydcyanhydrins (BE-Matrix):

	C	N	C	C	H	H	H
O	4	0	1	0	1	0	0
N	0	2	0	3	0	0	0
C	1	0	0	1	0	1	1
C	0	3	1	0	0	0	0
H	1	0	0	0	0	0	0
H	0	0	1	0	0	0	0
H	0	0	1	0	0	0	0

Tab. 2

Eine chemische Reaktion wird durch die Addition einer zweiten Matrix (M_R) zu der ersten (M_E) beschrieben, die neue Matrix (M_P) stellt das Bindungslektronenschema der neuen Produkte dar.

$$M_E + M_P = M_P$$

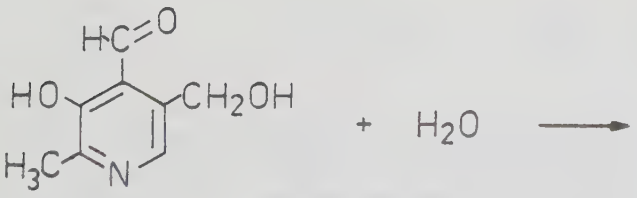
Durch dieses algebraische Modell der Chemie ist es möglich, ohne Darstellung von chemischen Strukturen chemische Sachverhalte zu verarbeiten.
Die Manipulationen verlaufen nach folgenden einschränkenden Regeln:

- Man beschränkt sich auf den Bereich der Elemente, die der Oktettregel gehorchen,
- man betrachtet nur die Valenzschale der Elemente. Die inneren Schalen spielen energetisch keine Rolle.
- Die Diagonalelemente der Matrix entsprechen den freien Elektronen eines Elements.
- Die Nicht-Diagonalelemente der Matrix entsprechen den Bindungselektronen in den kovalenten Bindungen.
- Die Summe der Eintragungen in der M_R muß = 0 sein, da Materie weder erzeugt noch vernichtet werden kann.

Die Vorgehensweise des Syntheseplanungsprogramms CICLOPS unterscheidet sich insofern von anderen Programmen, als eine Reihe von kleineren Molekülen dem isomeren Ensemble von Molekülen zugefügt und damit versucht wird, neue Produkte vorherzusagen. Auch die umgekehrte Vorgehensweise ist möglich: es wird versucht, ein gegebenes Zielmolekül in eine Reihe von kleineren Molekülen, z.B. H_2O, NH_3, HCN, CO_2, CO, Methanol oder Ethanol und einem größeren Molekül zu zerlegen. Es stehen sowohl der retrosynthetische Weg als auch die Methode der vorwärtsschreitenden Planung zur Verfügung.
Das Programm EROS benutzt das gleiche mathematische Modell der konstitutionellen Chemie wie CICLOPS. Es verfährt retrosynthetisch und gibt dem Benutzer keine Möglichkeiten zum

Eingriff in den Ablauf. Er kann nur entscheiden, ob er den Synthesebaum mehr in die Breite oder mehr in die Tiefe dirigieren möchte. Da das Programm Bindungen zwischen Atomen knüpft oder bricht, liegt es nahe, auch die dabei entstehenden Bildungsenthalpien zu berechnen. Man verwendet dazu keine quantenmechanische Näherungsrechnungen - dies wäre auch zu aufwendig - sondern man berechnet die Energien nach der Inkrement-Methode. Die erreichbare Genauigkeit beträgt 0,5 kcal/Mol. Die so errechneten Werte werden verwendet, um ein Maß für die Güte der gefundenen Vorschläge zu erhalten. Unter der Voraussetzung, daß die einfachen Elementarreaktionen thermodynamisch kontrolliert verlaufen, kommt man zu plausiblen Vorstellungen. Nur annähernd richtig werden die Voraussagen, wenn es darum geht, die Mengenverhältnisse von gleichzeitig entstehenden Produkten vorherzusagen.
Das Programm hat seine Leistungsfähigkeit z.B. bei der Berechnung der Valenztautomeren des Benzols, bei der Erzeugung der Isomere und Chlorierungsprodukte von n-Butan und bei der Lösung der Probleme bei der Synthese von β-Lactamen unter Beweis gestellt.

Beispiel: Biosynthese des Pyridoxals



Fortsetzung

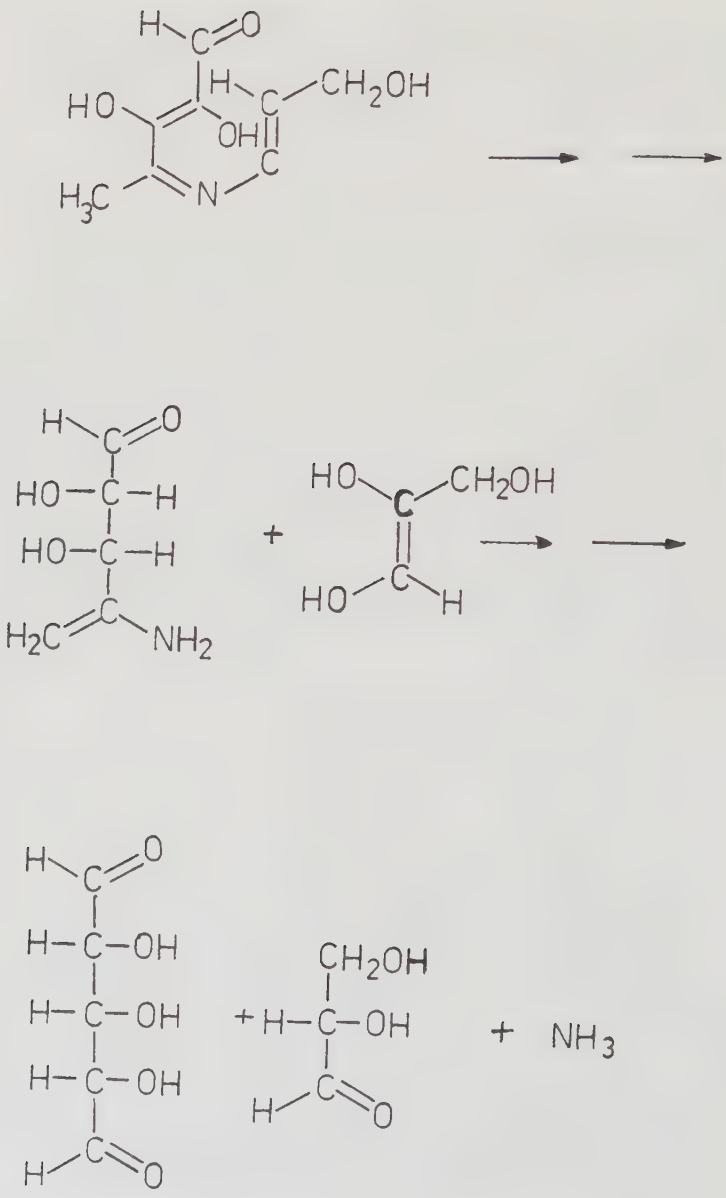


Abb. 12

Dieses mathematische Modell der Chemie kann neben der formalen Behandlung von chemischen Reaktionen und der Berechnung der Enthalpien von Reaktionsschritten auch da verwendet werden, um sich ein heuristisches Maß für die Kompliziertheit einer chemischen Reaktion zu definieren (76). Als Maß dafür dient die chemische Entfernung, die zwei Stoffe voneinander haben. Die Größe der BE-Matrix dient dazu, die Entfernung zwischen zwei Reaktionspartnern zu definieren. Dabei geht man von der Annahme aus, daß die Reaktion mit der geringeren Entfernung die günstigere sei. Das von Ugi formulierte Prinzip der minimalen chemischen Distanz (PMCD) erlaubt es, das Konzept der minimalen chemischen Strukturänderung rechnerisch nachzuvollziehen und Vergleiche zu treffen. Es lautet:

> Die meisten chemischen Reaktionen nehmen den Weg der minimalen chemischen Distanz, d.h. unter Verschiebung von möglichst wenig Valenzelektronen. Angewandt bei der Analyse von Synthesen komplizierter Naturstoffe wie Colchicin (Eschenmoser) und Strychnin (Woodward) zeigt es sich, daß das PMCD zur Abschätzung von Syntheseverläufen brauchbar ist.

Die Hauptbedeutung des PMCD besteht darin, daß man ein willkürfreies Auswahlkriterium hat, das in recht kurzer Rechenzeit Auskünfte über die vorgeschlagenen Synthesewege liefert. Es ist für ein Syntheseplanungsprogramm einfach, den Synthesebaum ins Unermeßliche wachsen zu lassen. Es existieren wenig Methoden zur Beschneidung des Baums. Die bisher praktizierten Regeln waren nicht willkürfrei und lieferten daher keine allgemeingültigen Ergebnisse. Das PMCD erlaubt eine Klassifikation der Vorschläge. Das Prinzip stellt eine heuristische Regel dar, die es erlaubt, Auswahlkriterien aufzustellen und Ähnlichkeitsmerkmale zu liefern. Neben der Syntheseplanung und der Berechnung von Enthalpiewerten und der Einführung des PMCD-Konzepts ist das Modell in der Lage, die Basis für ein Klassifikationssystem chemischer Reaktionen zu liefern.

Dazu ist es notwendig, für jede Reaktion eine eindeutige Notation zu entwickeln. Für spezielle Reaktionstypen hat ARENS (77 - 79) das Problem teilweise gelöst, allerdings beschränkt sich sein Ansatz auf wichtige Reaktionen, ausgehend von der Diels-Alder-Reaktion. Ein allgemeingültiges Notationssystem mit einer kanonischen Darstellung aller chemischen Reaktionen, das gleichzeitig die Basis für ein Dokumentations- und Retrievalsystem für chemische Reaktionen darstellt, befindet sich in der Entwicklung.

3. Aufgabenstellung und Ziele der Arbeit

3.1. Das Konzept der Direktspeicherung von Sachverhalten

Die Direktspeicherung von Sachverhalten führt den Benutzer ohne Umwege über die Literaturhinweise zu den als für die Fragestellung relevant erkannten Sachverhalten.
Eine Voraussetzung der Konzeption der Speicherformate für Daten und Fakten ist die Systemanalyse, die die besonderen Wissensstrukturen, die Dynamik und die Problemlösungsmethoden der Chemiker berücksichtigt.
Als Chemiegebiet wurde die Chemie der Chlorformamidine ausgewählt.

Das Konzept des direktspeichernden Informationssystems gliedert sich in drei Teile:

1. Grundlage ist ein Fakten- und Datenspeichersystem.
   Zu jeder Verbindung im System werden folgende Daten - soweit sie vorhanden und zugänglich sind - nachgewiesen:
   - Namen
   - Summenformel
   - Schmelzpunkt mit Erläuterungen

- Siedepunkt mit Erläuterungen
- IR-Spektrum mit Interpretationen
- ^{1}H-NMR-Spektrum mit Peakzuordnungen
  und
- pK_b-Werte

Derivate und Vorgänger einer Verbindung sind über die Beschreibung der Reaktionsgleichung, die auch die Literaturhinweise enthält, miteinander verknüpft.

2. Das zweite Teilsystem erlaubt die Abfrage der gespeicherten Daten in einer benutzerfreundlichen Sprache und die Durchführung einfacher mathematischer Operationen.

3. In diesem Teil erlaubt das Abfragesystem die Durchführung von logischen Operationen und Berechnungen, die auf heuristischen Regeln basieren. Neue Beziehungen zwischen Daten und Fakten werden von System selbständig generiert.

Die heuristischen Regeln werden während der Systemanalyse identifiziert, soweit als möglich formalisiert und im Programm implementiert.

3.2. Vorstellung der allgemeinen Konzeption von CHEMIS

Klassifikationen, die für eine Systemplanung als Ordnungsprinzip des Wissens als gültig angenommen wurden, werden durch ihre starre Aspektbezogenheit im Laufe der Zeit immer unzulänglicher und rigider, da sie die dynamische Wissensentwicklung nicht mehr adäquat abbilden können. Durch die Direktspeicherung und die Abfragemöglichkeiten strebt man an, das System von den Beschränkungen der herkömmlichen Ansätze frei zu halten. Bisherige Untersuchungen zum Informationsverhalten von Chemikern (2, 80)

haben sich vorwiegend damit befaßt herauszufinden, wie die Chemiker sich mit Hilfe der ihnen zur Verfügung stehenden Literaturinformationssysteme informieren.
Dabei wurde ermittelt,

- welche Methoden zur laufenden Informationsbeschaffung dienen,
- welche Mittel benutzt werden, um sich bei neu auftauchenden Fragen einen Überblick zu verschaffen,
- welche Dokumentationsdienste benutzt werden und
- welche Anforderungen an die Dienstleistungen gestellt und wie diese erfüllt werden.

Aufgrund der dabei erhaltenen Aussagen werden Verbesserungen am Gesamtsystem vorgenommen, die die spezifischen Bedürfnisse des einzelnen Chemikers zu wenig berücksichtigen. Durch das hier gewählte Vorgehen der partizipatorischen Systementwicklung wird die Auseinanderentwicklung von Benutzerbedürfnissen und Systemleistung verhindert. Es ist charakteristisch für den hier verfolgten Untersuchungsansatz, daß der potentielle Benutzer möglichst früh in die Systementwicklung mit einbezogen wird. Er kann die unterschiedlichen Stadien des Prozesses laufend mit nachvollziehen und überprüfen, welche Konsequenzen seine Vorschläge auf das praktische Rechercheergebnis haben. Die Bereitschaft der Benutzer, sich aktiv am Systementwicklungsprozeß zu beteiligen, wächst. KUNZ und RITTEL stellen daher fest, daß von der Verbesserung der Dokumentationsmethodik alleine nicht der wesentliche Fortschritt in der Informationsversorgung zu erwarten ist (81):

> "Mit den Methoden der Systemanalyse lassen sich charakteristische Denkprozesse von Systemlösern, auch des Chemikers beschreiben, woraus Schlüsse über die Struktur seines Wissens, dessen Veränderungen und seiner Nutzung gezogen werden können. Sie liefern Hinweise für den Entwurf besser organisierter "äußerer" Informationsspeicher, welche den Chemiker bei der Lösung seiner

Aufgabe unterstützen. Es hat sich gezeigt, daß Informationssysteme keine separablen Teilsysteme des Forschungssystems sind: welche Daten man speichert und auf welche Weise, hängt von den Zwecken ab, für die sie benutzt werden sollen; zudem beeinflussen die Praktiken der Datenverarbeitung und -speicherung nicht nur die Wirksamkeit der Forschung, sondern auch ihr Vorgehen."

Ziele der Arbeit sind:

- Systemanalyse des Problemlösungsverhaltens von organischen Chemikern am Beispiel der Chlorformamidin-Chemie,
- Identifikation von Informationsbedürfnissen und typischen Fragestellungen,
- Aufbau eines direktspeichernden Informationssystems, für präparativ arbeitende Chemiker,
- Entwicklung einer benutzerfreundlichen Abfragesprache zur Formulierung der Suchfragen
- Identifizierung, Formalisierung und Implementierung von heuristischen Regeln zur Strukturierung des vorhandenen Wissens, zur Generierung neuer Beziehungen zwischen Daten und Fakten und zur Klassenbildung nach vom Benutzer vorgegebenen Regeln.

4. Beschreibung der Systementwicklung

4.1. Vorarbeiten

Der hier verfolgte Ansatz wurde schon von RAMI (2) auf dem Gebiet der Peptidchemie erprobt. Diese Fallstudie bezog sich auf einen relativ homogenen Kreis von etwa 10 Diplomanden und Doktoranden, die Themen aus dem Bereich der Konformationsanalyse von Peptiden mit Hilfe der ^{1}H- und ^{13}C - NMR-Spektroskopie untersuchten.

Die dabei angewandte Vorgehensweise besteht aus folgenden Teilschritten:

- Ermittlung von Fragestellungen,
- Herausarbeiten von typischen Anfragestrukturen,
- Ableitung von Entwurfskriterien
- Aufstellung von strukturierten Listen,
- Beschreibung von individuellen Problemlösungsvorgängen in einem Flußdiagramm,
- Formulierung von allgemeinen Problemlösungsvorgängen in Flußdiagrammen,
- Verwendung von Heuristiken bei der Strukturierung der Programme

Das von Rami ausgewählte Gebiet der Peptidchemie hat den Vorteil, daß es relativ eng begrenzt ist und daher mit relativ wenig Aufwand eine gute Modellierung des Wissens erreicht werden kann: es gibt eine bestimmte Anzahl von Aminosäuren, die sich für diese Untersuchungen eignen, es liegen dadurch auch die notwendigen Schutzgruppen für die Amino- und Carboxylgruppe fest, dazu kommen noch einige Kupplungsmethoden für die geschützten und aktivierten Aminosäuren und die Abspaltungsmethoden bei den Oligopeptiden. Diese stoffliche und methodische Vielfalt wird von Strategien zur Peptidsynthese gegliedert.

4.2. Zur Auswahl des behandelten Gebiets

Zur Durchführung der Arbeit wurde das Gebiet der Chemie der Chlorformamidine, genauer das der N^2-Arylchlorformamidine ausgewählt.
Für die Auswahl war wichtig, daß die Mitarbeiter des Arbeitskreises persönlich bekannt sind. Die Befragung wird einfacher, die Ergebnisbewertung wird präziser und die Antwortbereitschaft wächst.

Vereinfachtes Schema der Peptid-Synthesen

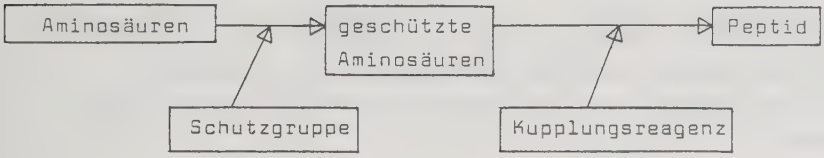


Abb. 13

Vereinfachtes Schema der Chlorformamidin-Chemie

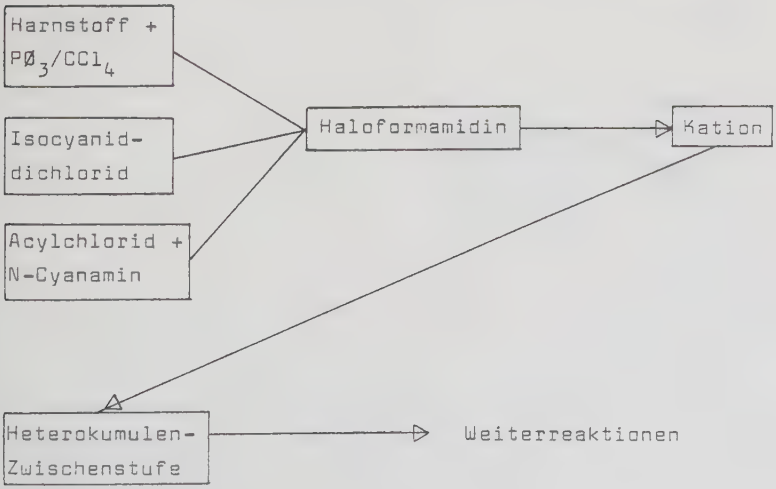


Abb. 14

Die Chemie der Chlorformamidine ist nicht so übersichtlich aufgebaut wie die der Aminosäuren und Peptide. Während bei der Peptid-Chemie das Syntheseziel, das aufzubauende Peptid, festliegt und ausgearbeitete Strategien den Weg zum Ziel ebnen, ist die Chlorformamidin-Chemie weniger gradlinig, es können auf jeder Synthesestufe Nebenprodukte auftreten, je nach eingesetztem Nukleophil ergeben sich unterschiedliche Produkte mit verschiedenen Weiterreaktionsmöglichkeiten. Chlorformamidine unterscheiden sich drastisch in ihrer Reaktionsfähigkeit: so gibt es Stoffe, die mittels eines Gemisches aus Dioxan/H_2O unter KOH-Katalyse hydrolisiert werden müssen, andere wiederum zerfließen beim Stehen an der Luft. (Eine genaue Diskussion der Hydrolyseempfindlichkeit erfolgt bei den Heuristiken für diese Eigenschaften, s.S.142)
Ein Mechanismus steht bei der Chlorformamidin-Chemie im Vodergrund: Die Dissoziation in ein Cl^- und in ein Amidinium-Kation:

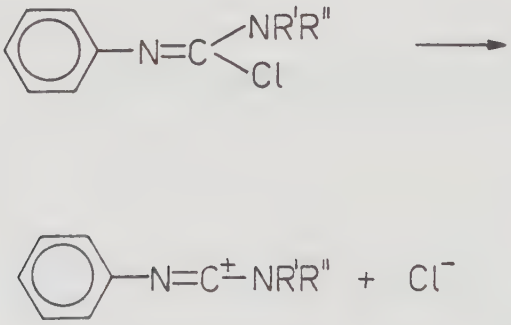
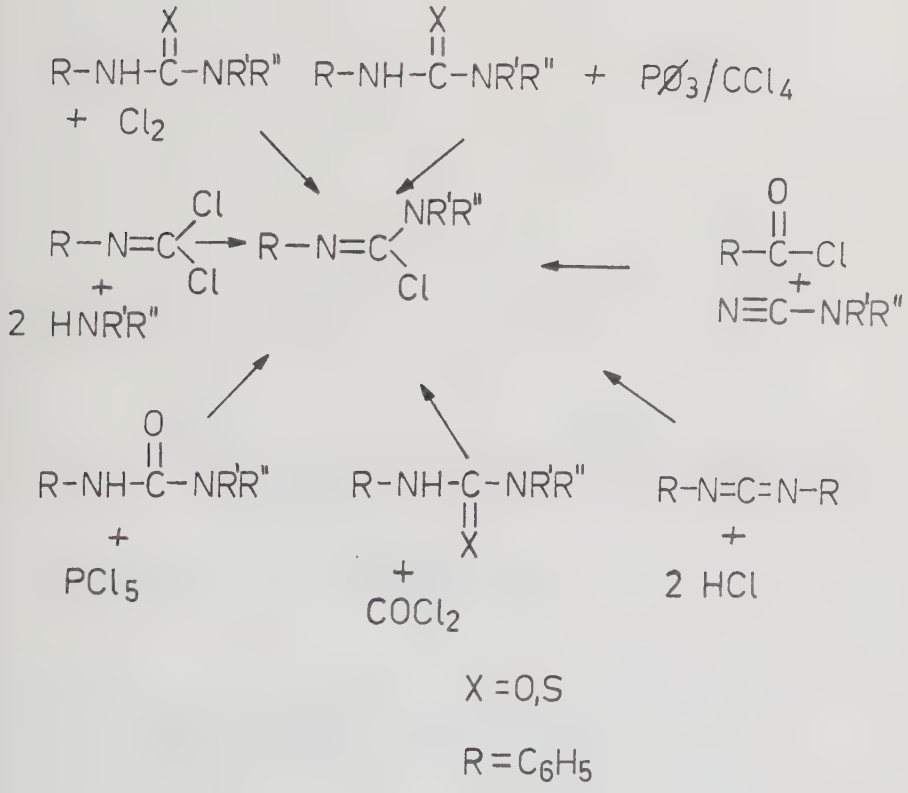


Abb. 15

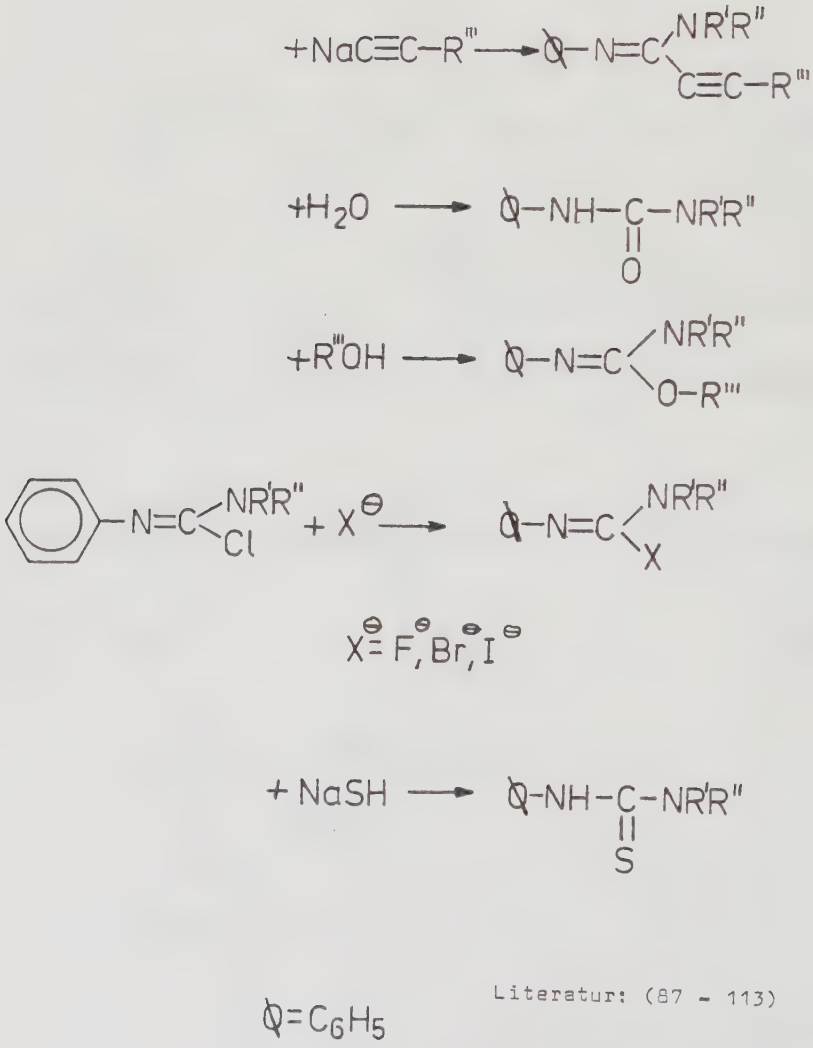
Überblick über die Chemie der Chlorformamidine:

1. Synthesemöglichkeiten:



Fortsetzung

2. Reaktionen der Chlorformamidine:



Literatur: (87 - 113)

Abb. 16

4.3. Systemanalyse des Forschungsprozesses präparativ arbeitender Chemiker

Die Methode der Systemanalyse in der Forschung beruht auf der Hypothese, daß sich das Denken und Arbeiten eines Chemikers mittels eines Modells des Problemlösungsverhaltens befriedigend beschreiben lassen. Dabei ist wesentlich, daß Antworten, Fragen und die bei der Arbeit benutzten Informationsmittel nicht isoliert gesehen werden dürfen, sondern als Einheit aufzufassen sind, um der Gefahr der Suboptimierung zu entgehen.
Jeder Forscher hat spezifische Informationsbedürfnisse. Selbst für Chemiker, die auf dem gleichen Gebiet arbeiten, spielen unterschiedliche Vorkenntnisse, alte Informationsgewohnheiten und verschiedene Spezialbedürfnisse eine ausschlaggebende Rolle. Dennoch ist es unmöglich, für jeden Chemiker ein persönliches Informationssystem zu realisieren. Deshalb wird versucht, durch die Systemanalyse gruppenspezifische Informationsbedürfnisse einiger Chemiker herauszufinden, diese dienen als Grundlagen und Bestimmungsgrößen für die Entwicklung eines Informationssystems.
Dabei war zu prüfen, inwieweit die von Rami vorgeschlagene Struktur des Peptid-Informationsystems als Basis für das Gebiet der Chlorformamidin-Chemie verwendet werden kann.

Man steht in der Chemie ab und zu vor der Tatsache, daß sich eine Reaktion nicht wie geplant durchführen läßt. Der Chemiker steht nun vor der Aufgabe, einen Ausweg zu finden. Er muß die Bedingungen variieren. Es stehen ihm folgende Möglichkeiten zur Auswahl:

- die Reaktionsbedingungen
  - Lösungsmittel
  - Katalysatoren
  - Druck
  - Temperatur
  - Konzentrationen und Reaktionspartner
  - Reaktionsdauer

und
- die Edukte.

In welcher Richtung soll er nun einen Ausweg suchen?
Das Verfahren der Systemanalyse gliedert sich in dieser Arbeit in folgende Teilschritte:
- Problemformulierung:
  Darunter versteht man die ausführliche verbale Formulierung der Frage, so wie sie sich aus dem Zusammenhang ergibt. Einschränkungen bei der Formulierung sollten vermieden werden.
- Hypothesenbildung:
  Zur Lösung eines Problems bedient man sich eines Lösungsansatzes. Bei der Formulierung der Hypothese wird der Informationsbedarf identifiziert und die Informationsquellen beschrieben.
- Identifikation und Bewertung der Quellen:
  Die Quellen liefern die Informationen, die zum Aufbau einer Hypothese und damit zum Lösungsversuch eines Problems benötigt werden. Es sind zwei Fälle zu unterscheiden.
  - Die benötigte Quelle existiert, d.h. die Frage kann mit den existierenden Informationsmitteln beantwortet werden. Die Verbesserung dieses Suchprozesses, die Zugänglichkeit der Quellen und die eventuell notwendige Beschaffung der Originalarbeiten sind nicht Gegenstand dieser Arbeit.
  - Die erforderliche Quelle existiert nicht. Der Informationsbedarf stößt ins Leere. Liegt der Fall vor, so hat man die Aufgabe, die Forderungen an die Datensammlung zu spezifizieren, so daß eine Modelldatensammlung aufgebaut werden kann, die in der Lage ist, exemplarisch aufzuzeigen, wie die Bedürfnisse erfüllt werden können. Inhalt dieser Datensammlungen sind Daten, Sachverhalte und ausgewählte Ausschnitte aus dem Hintergrundwissen, das zur Lösung eines Problems erforderlich ist.

Liefern die Quellen keine Daten und ist es auch nicht abzusehen, daß durch Organisation des vorhandenen Wissens eine Antwort gefunden werden kann, so sind Experimente zu planen. Dabei kann ein Informationssystem auf zweierlei Arten Hilfestellung leisten. Es kann vorhandene Daten nachweisen und dadurch Experimente, die immer Aufwand erfordern, überflüssig machen, oder es kann die Rahmenbedingungen, unter denen die Experimente geplant werden, verbessern. Die experimentellen Ergebnisse gewinnen dadurch an Aussagekraft. Ergebnisse können sowohl als numerische Daten als auch in Form von Sachverhalten vorliegen (Schmelzpunkte, Spektraldaten und Beschreibung des reaktiven Verhaltens).

Da aber im allgemeinen die Vollständigkeit der Daten nicht gewährleistet ist und die Frage nach allen formal möglichen Einzelheiten nicht notwendig ist, ist es sinnvoll, eine Datenauswahl zu treffen.

Die dabei angewandte Vorgehensweise stellt sich so dar:
Ein Problem wird auf alle Parameter untersucht, die zu seiner Beschreibung zur Verfügung stehen. Es wird geprüft, welche Parameter zur Lösung des Problems in erster Näherung nicht unbedingt notwendig sind, sie werden vernachlässigt. Mit den übrig gebliebenen Daten versucht man, eine Näherungslösung zu erreichen. Hat man eine Näherungslösung skizziert, so kann man beginnen, die Lösung zu verbessern, indem man weitere Beschreibungselemente in die Lösungsansätze mit einbezieht.
Der Lösungsansatz besteht nun darin, die als wesentlich erkannten Parameter zu variieren. Auf der Basis der entwickelten Hypothese zur Problemlösung werden die Konsequenzen untersucht, die die Änderung einer Größe auf die Lösung des Problems hat.
Ein Syntheseplanungsunterstützungssystem für das Gebiet der Chlorformamidine sollte in der Lage sein, den jeweiligen Problemlösungsprozeß zu unterstützen.

Er kann, z.B. von einer bekannten Reaktion ausgehend, sich Reaktionen nachweisen lassen, die sich von ihr in einigen vorgebenen Aspekten möglichst wenig unterscheiden.
Er kann fragen, welche von denen im Speicher vorhandenen Reaktionen seiner Ausgangsreaktion hinsichtlich der Edukte, der Produkte oder der physikalischen Beschreibung der Reaktionsbedingungen möglichst ähnlich sind.

Für Stoffe existieren verschiedene Daten, die als Ähnlichkeitsmerkmal herangezogen werden können:

- Summenformel
- Schmelzpunkt
- Siedepunkt
- Brechungsindex
- IR-Spektrum (einzelne Banden oder eine Bandengruppe)
- NMR-Spektrum
- spezifische Drehung
- UV-Spektrum

  etc.

Solche Ähnlichkeitsrelationen lassen sich als zweistellige Beziehungen und Funktionen, die Eigenschaften beschreiben, die zwei Objekten gemeinsam sind, verstehen.
Sie bestenen aus drei Elementen:
- dem Grundobjekt,
- dem zu vergleichenden Objekt und
- der Relation, die die Verbindung zwischen den beiden Objekten herstellt.

Die Anwendung der Ähnlichkeitsrelation in CHEMIS stützt sich auf die Beschreibung der Reaktionsfähigkeit der Stoffe. Durch die direkte Beschreibung des Verhaltens der Stoffe werden die Probleme der Strukturdarstellung und Strukturverschlüsselung umgangen. Die Verarbeitung der chemischen Sachverhalte führt direkt zu den relevanten Antworten.
Strukturelle und physikalische Ähnlichkeiten sind dabei die Kriterien bei der Beurteilung chemischer Verbindungen hinsichtlich ihrer Ähnlichkeit.
Bei der Untersuchung des Ähnlichkeitsverhaltens der Stoffe geht man von der Annahme aus, daß eine Substanz S_x einer vorgegebenen Substanz S_1 um so ähnlicher ist, je mehr die Reaktionsbeschreibungen der beiden Stoffe übereinstimmen.

Dieses ähnliche Reaktionsverhalten läßt folgende Fallunterscheidungen zu:
1. Man untersucht die Anzahl der beiden Stoffe gemeinsamen Folgeprodukte (Derivat-Ähnlichkeit).
2. Man untersucht die Anzahl von Edukten, die beiden Stoffen gemeinsam sind (Ursprungs-Ähnlichkeit).
3. Man untersucht die Anzahl der beiden Stoffe gemeinsamen Reaktionspartner (Substitutions-Ähnlichkeit).

Die Formalisierung und Programmierung der Kriterien führt zu Listen, in denen die Stoffe nach fallender Ähnlichkeit geordnet sind.

Liste der zu Ethanol substitutiv ähnlichen Stoffe:

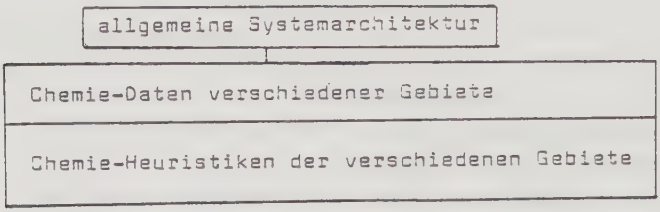
(Fall 3)

```
sae 1135
SUBSTITUTIVE AEHNLICHKEIT:
VERB.1   VERB.2   ANZ.1   ANZ.2   ANZ.G.     % 1      % 2
1135     1135       68      68       68   100,00   100,00
1135     1102       68      47       38    55,88    80,85
1135     1378       68      46       38    55,88    82,60
1135       15       68      57       30    44,11    52,63
1135     1189       68      30       28    41,17    93,33
1135     1162       68      34       27    39,70    79,41
1135     1397       68      35       27    39,70    77,14
1135     1416       68      30       27    39,70    90,00
1135     1146       68      29       26    38,23    89,65
1135     1181       68      35       26    38,23    74,28
1135     1188       68      52       26    38,23    50,00
1135        3       68     124       25    36,76    20,16
1135     1244       68      26       24    35,29    92,30
1135     1247       68      30       24    35,29    80,00
1135     1101       68      43       21    30,88    48,83
1135     1104       68      26       21    30,88    80,76
1135     1137       68      27       21    30,88    77,77
1135     1139       68      43       21    30,88    48,83
1135     1204       68      44       21    30,88    47,72
AUSGABE FORTSETZEN?  (J;N)
j
1135     1227       68      26       21    30,88    80,76
1135     1123       68      26       20    29,41    76,92
1135     1236       68      23       20    29,41    86,95
1135     1149       68      29       19    27,94    65,51
1135     1399       68      41       19    27,94    46,34
1135     1148       68      21       18    26,47    85,71
1135     1179       68      33       18    26,47    54,54
1135     1380       68      26       18    26,47    69,23
1135        1       68      30       16    23,52    53,33
1135        2       68      39       16    23,52    41,02
1135        5       68      68       16    23,52    23,52
1135       21       68      21       16    23,52    76,19
1135     1277       68      20       16    23,52    80,00
1135     1170       68      17       15    22,05    88,23
1135     1232       68      17       15    22,05    88,23
1135       39       68      16       12    17,64    75,00
1135     1103       68      22       12    17,64    54,54
1135        8       68      44       11    16,17    25,00
1135     1186       68      15       11    16,17    73,33
1135     1194       68      13       11    16,17    84,61
AUSGABE FORTSETZEN?  (J;N)
n
```

.
.
.

Die zahlenmäßige Exaktheit der Werte darf nicht darüber hinwegtäuschen, daß die Werte auf rein statistischer Aus-- wertung der Daten beruhen. Die Zahlen spiegeln den augenblicklichen Speicherinhalt des Systems wider und erlauben nur Aussagen, die sich in diesem Rahmen bewegen. Insbesondere soll offen bleiben, ab welchem Zahlenwert eine Ähnlichkeitsaussage nicht mehr sinnvoll sein soll bzw. ob man nicht zur Erhöhung des Aussagewerts den Mittelwert aus allen drei Ähnlichkeitsberechnungen verwenden sollte.
Eine weitere Einschränkung ist in der überproportionalen Berücksichtigung der Chemie der Aminosäuren oder der Chlorformamidine, Harnstoffe und Isocyaniddichloride zu sehen.
Die Einschränkung auf das Gebiet der Arbeit bedeutet, daß die Ähnlichkeitslisten als vorläufiges Versuchsprotokoll zu betrachten sind, deren Inhalt sich mit dem weiteren Aufbau des Systems ändern wird.
Bei diesen Vergleichswerten wurde darauf geachtet, daß jeder Chemiker in der Lage sein sollte, sich seinen Auschnitt aus der Chemie mittels den vorher diskutierten Ähnlichkeitsmerkmalen zu organiseren und sich Verbindungen vom System als ähnlich zusammenführen zu lassen, die sich in seiner Chemie als ähnlich erwiesen haben.

Prinzipiell bieten sich zwei alternative Systemkonzepte an:



oder

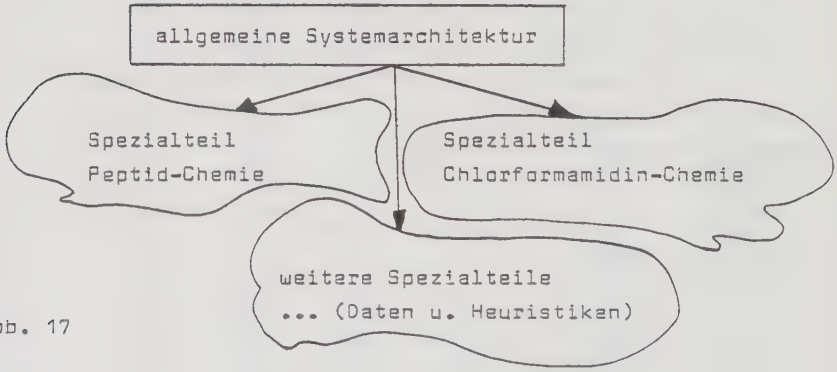


Abb. 17

Die Untersuchungen beider Gebiete zeigten, daß in der gegenwärtigen Entwicklungs- und Ausbauphase des Systems die Trennung der Systeme nicht zweckmäßig ist, da beide Gebiete sich in ihrer Datenstruktur noch ähnlich genug sind, um die Behandlung in einem gemeinsamen System zu rechtfertigen.

Die Systemanalyse eines ausgewählten Gebiets der organischen Chemie hat zum Ziel, aus der Gesamtheit der zur Verfügung stehenden Informationen diejenigen herauszufinden, die für den Forschungsprozeß wesentlich sind, d.h. die typischen Anfragen des präparativen organischen Chemikers bei der Lösung konkreter Probleme zu identifizieren. Die Ergebnisse dieser Systemanalyse dienen als Grundlage für die Formalisierung dieses Fragetypus.

Beispiel: Um nach der Bredereck-Methode Chlorformamidine zu synthetisieren, benötigt man Carbonsäurehalogenide als ein Edukt. Die Auswahl der zur Verfügung stehenden Acylhalogenide wird durch die Tatsache eingeschränkt, daß sich die Reaktion nur dann mit befriedigenden Ausbeuten durchführen läßt, wenn auf die Verwendung eines Lösungsmittels verzichtet wird. Die Reaktion wird durch Zusammenschmelzen der Komponenten durchgeführt (84, 85).
Die Reaktionstemperatur beträgt 70 °C - 80 °C. Damit ist ein zweites Merkmal der Stoffklasse festgelegt: Sie sollten etwa in diesem Teperaturintervall schmelzen, um zusammen mit der zweiten Komponente, dem N-Cyan-Amin ein einheitliches Gemisch ergeben zu können. (Das N-Cyan-Amin muß frisch destilliert werden, bevor es eingesetzt wird, sonst wird der Reaktionsansatz schwarz und zersetzt sich heftig.) Als weitere einschränkende Bedingung kommt hinzu, daß das Acylhalogenid ziehend substituiert sein muß. Aromatische Säurechloride lassen die Variation dieses Prinzips zu.

Die Frage lautet also: Welche aromatischen Carbonsäurehalogenide sind bekannt, die am Kern ziehend substituiert sind und die in einem Temperaturbereich zwischen 60 °C und 90 °C schmelzen?

Man fragt gleichzeitig nach:

- strukturellen Merkmalen
  - aromatische Verbindung
  - Carbonsäurederivat
  - Substitution am aromatischen Ring
- physikalische Daten
  - Schmelzbereich
- Reaktivitätsdaten
  - elektronenziehende Substitution am aromatischen Ring, oder allgemein: eine Aktivierung der Bindung zum Halogen-Atom
  - Reaktionen in der Schmelze

Keines der bisherigen Informationssysteme ist in der Lage, diese Frage zu beantworten. Allenfalls Teilantworten sind zu erhalten. Mit den neuen Diensten "CAS-online" und "DARC" sind Recherchen nach chemischen Strukturen möglich, bei "DARC" ist die Kopplung zwischen strukturellen Daten und den Ergebnissen einer Recherche in konventionellen bibliographischen Datenbasen einfach möglich.
Literaturinformationssysteme können von ihrer Konzeption her diese Frage nicht beantworten. Strukturelle Merkmale zählen im allgemeinen nicht zum Vokabular ihrer Dokumentationssprache. Darüberhinaus fehlt ihnen der Datenteil.
Dateninformationssysteme weisen Daten zwar nach, anschließend geben sie Hinweise auf die Referenzliteratur.
Konventionelle Systeme können also nur isolierte Fragen beantworten. In der Forschung treten solche Fragen selten auf. Die Übersetzung von Fragen in die Sprache des

Dokumentationssystems ist nicht selten ein verlustreicher Prozeß. Die Frage muß der Leistungsfähigkeit der benutzten Informationssysteme angepaßt sein. (Beispiele für diesen Prozeß finden sich in den "CAS Printed Access Tools" (45).)

Die Informationssysteme erweisen sich als nicht besonders innovationsfreundlich, da sie nicht schnell genug auf wechselnde Gegebenheiten ansprechen können.
Aus dieser Situation ergibt sich die Forderung, sich nicht zu früh auf eine Systematik für das Informationssystem festzulegen. Die Systemkonzeption sollte so flexibel angelegt sein, daß sich während der Arbeit ergebende Gesichtspunkte mit einbeziehen lassen.
Festlegungen in der Datenstruktur, die aus praktischen Gründen getroffen werden müssen, sollten sich ohne großen Aufwand ändern lassen, wenn die Anforderungen an das System sich ändern.

Die Systemanalyse des Forschungsprozesses hat als verallgemeinerbare Randbedingungen ergeben:

- Die Größe der Reaktionsansätze liegt durchweg im mg- bis g-Bereich. Bei Vorprodukten wird diese Größenordnung überschritten.
- Die Reaktionsführung ist diskontinuierlich. Es ist daher nicht notwendig, die in der industriellen Praxis notwendigen Parameter wie z.B. Durchflußmenge, Umsatz oder Viskosität zu messen oder zu regeln. Es entfallen die Heiz- und Kühlprobleme, da die erforderliche Reaktionstemperatur im Labor schnell durch Eintauchen des Reaktionsgefäßes in ein Medium mit der gewünschten Temperatur erreicht wird.
- Die Durchmischung des Reaktionsmediums bereits ebenfalls keine besonderen Probleme, ein Magnetrührer ist ausreichend.
- Extreme Bedingungen, z.B. sehr hohe oder sehr niedrige Temperaturen und hohe Drucke sind mit den im Labor zur Verfügung stehenden Mitteln schwer zu realisieren.

Sie bedeuten einen Zeitverzug bei der Arbeit und werden daher nach Möglichkeit vermieden.
Der im Labor gängige Temperaturbereich erstreckt sich von

- ca. - 30 °C (Temperatur einer Eis/Kochsalz-Kältemischung)

bis zu

- ca. + 150 °C.

Reaktionen, die in Lösung nicht bis zu dieser Temperatur mit befriedigender Geschwindigkeit oder mit zufriedenstellender Ausbeute ablaufen, bedürfen oft der Unterstützung eines Katalysators.
Eine weitere Temperaturerhöhung hat meistens die Zunahme der Zersetzungsreaktionen zur Folge.
Erhöhter Druck zählt ebenfalls zu den Bedingungen, die im Labor nicht so ohne weiteres zu realisieren sind. Reaktionen, die im Bombenrohr oder im Autoklaven durchgeführt werden müssen, erfordern einen erhöhten Sicherheitsaufwand.

Je höher also Reaktionstemperatur und Druck sind, desto schwieriger werden Reaktionsvorbereitung und -führung.

Es zeichnet sich eine Grenze für die praktische Durchführbarkeit von Reaktionen im Labor ab:
Sobald die erforderliche Temperatur die Höhe von 150 °C über- oder - 30 °C unterschreitet und der anzuwendende Druck den Atmosphärendruck übersteigt, sind umfangreichere Vorbereitungen erforderlich.

Die Untersuchungen haben ergeben, daß das CHEMIS-System "Chlorformamidine" sich in zwei Reaktionskategorien einteilen läßt:

- in Reaktionen, die unter milden Bedingungen ablaufen und
- in Reaktionen, die drastische Bedingungen benötigen.

In dieser Arbeit werden daher milde Bedingungen definiert, wenn die Reaktionstemperatur sich im Bereich zwischen - 30 °C und + 150 °C bewegt und der Druck gleich dem Atmosphärendruck ist.
Drastische Bedingungen liegen dann vor, wenn Druck oder Temperatur die oben definierten Bereiche verlassen.

Die Einteilung ist den Bedürfnissen der Chlorformamidin-Chemiker angepaßt. Ein anderer Benutzer ist ohne weiteres in der Lage, die Klassifikation der Reaktionsbedingungen nach seinen speziellen Bedürfnissen vorzunehmen oder überhaupt auf eine Sortierung nach den Merkmalen zu verzichten.
Zusammen mit anderen Gesichtspunkten, wie z.B. der erzielten Ausbeute oder der Verwendung eines bestimmtem Lösungsmittels, kann man versuchen, sich ein Gesamtbewertungskriterium für die Eignung einer Reaktion vom System bilden zu lassen.

Um einer alternativen Reaktion vorgezogen werden zu können, sollte eine Reaktion folgende Bedingungen erfüllen:

- Sie sollte unter milden Bedingungen verlaufen.
- Sie sollte die größere Ausbeute haben.
- ein bestimmtes Lösungsmittel sollte verwendet werden.

Die Verwendung von wasserfreien Lösungsmittel zählt nicht zu den Kriterien, die eine Reaktion wesentlich erschweren.

Wenn der Wasserausschluß eine entscheidende Bedingung für das Gelingen einer Rekation ist, so ist es notwendig, in einer Kommentarzeile bei der Beschreibung der Reaktionsbedingungen auf die Notwendigkeit des wasserfreien Arbeitens hinzuweisen.
Bei der Aufstellung der Liste der hydrolyseempfindlichen Chlorformamidine wird dem Rechnung getragen. (s.S. 149)

Die Reaktionszeit spielt im Labormaßstab keine überragende Rolle. Die meisten in Frage kommenden Reaktionen sind innerhalb eines Tages beendet. Daher wird in der jetzigen Ausbaustufe von CHEMIS auf den Nachweis von Reaktionszeiten und ihre Einbeziehung in die Bewertung von Reaktionsbedingungen verzichtet. Ein weiterer Gesichtspunkt ist, daß die Angaben über die Reaktionszeiten in den Arbeitsvorschriften häufig fehlen.
Als weiteres wichtiges Merkmal bei der Beschreibung einer Reaktion spielt die erzielte Ausbeute eine Rolle. Je nach der Bedeutung einer Reakton im Syntheseablauf hat die Ausbeute einen unterschiedlichen Stellenwert. Je weiter fortgeschritten im Syntheseablauf das fragliche Produkt benötigt wird, desto wesentlicher wird die Frage nach der Ausbeute. Ist das Produkt aus anderen Gründen wichtig,z.B. als Belegsubstanz für einen vorgeschlagenen Mechanismus, so spielen Ausbeutegesichtspunkte nur eine geringere Rolle, solange sichergestellt werden kann, daß für die Analysen genügend Substanz zur Verfügung steht. Zum wesentlichen Kriterium wird die Ausbeute bei der Beurteilung von mehrstufigen Reaktionsfolgen. Im System werden sowohl die Ausbeute bzw. der Ausbeutebereich als auch die Gesamtausbeute für jeden einzelnen Schritt nachgewiesen. Dadurch ist es möglich, genau zu verfolgen, bei welcher Einzelreaktion der kritische Punkt bei der Ausbeute liegt, so daß es möglich ist, gezielte Alternativen für ungünstige Reaktionen zu suchen.

Bei der Berechnung der Ausbeute kann man zwei verschiedene Wege gehen (86):

Schema der konvergierenden Strategie:

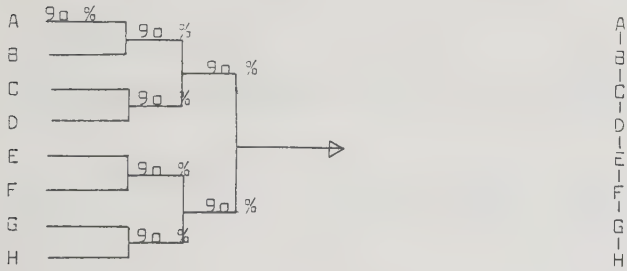


Abb. 18

Schema der sequentiellen Strategie:

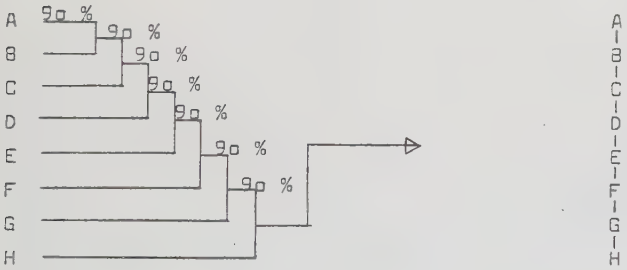


Abb. 19

Die Einzelausbeute jeder einzelnen Reaktion sei in diesem Beispiel 90 %.
Bei der sequentiellen Synthese ergibt sich für das lineare Peptid aus den Aminosäuren A - H eine Gesamtausbeute von 47,8 %, bei der konvergierenden Synthese dagegen erreicht man eine Gesamtausbeute von 73 %. Mit der Anzahl der Stufen steigt der Vorsprung der konvergierenden Synthese.

Lösungsmittel

Das Lösungsmittel spielt in der Diskussion der Reaktionsbedingungen für Chlorformamidine, die zu variieren sind, um eine geplante Synthese durchführen zu können, eine wichtige Rolle.

Liste der für Lösungsmittel wichtigen Eigenschaften (Auswahl):

- Schmelzpunkt
- Siedepunkt
- Flammpunkt
- Zündtemperatur
- Dichte
- Dipolmoment
- Dielektrizitätskonstante
- Leitfähigkeit
- Viskosität
- Vergleichsspektren
    - IR-Spektrum
    - NMR-Spektrum
    - UV-Spektrum
    - Massenspektrum
    - HPLC-Vergleichsdaten
- Bezugsquellen
- Reinigungs- und Absolutierungsvorschriften
- Preise

Absolutierungsschema für Lösungsmittel:

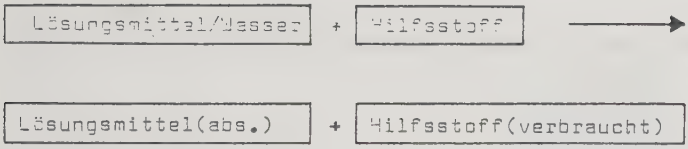


Abb. 20

Rein formal betrachtet findet eine chemische Reaktion statt, Wasser wird durch die Umsetzung selektiv aus dem Gemisch entfernt.

Die Information, ob für eine bestimmte Reaktion Wasserausschluß notwendig ist oder nicht, erfährt der Anfrager in CHEMIS aus der Kommentarzeile bei der Reaktionsbeschreibung.

Zur Erkennung der Lösungsmittel bedient man sich des folgenden Verfahrens:
In der Beschreibung einer Reaktion in CHEMIS existiert die Kategorie "Lösungsmittel". Man definiert sich eine Vorschrift, die alle Stoffe, die mindestens einmal als solches verwendet wurden, in eine eigene Liste überführt. Diese Liste stellt sozusagen eine Momentaufnahme des Systemzustands dar.

Liste der vom System abgespeicherten Lösungsmittel:

```
LOESUNGSMITTEL IN DEN REAKTIONEN
    3:  WASSER
   12:  SCHWEFELSAEURE
   32:  THIONYLCHLORID
 1102:  METHYLALKOHOL
 1147:  DIOXAN
 1169:  BENZOL
 1181:  ESSIGSAEURE
 1202:  DIAETHYLAETHER
 1236:  ACETANHYDRID
 1315:  BUTANDIOL-(1.4)
 1607:  TRIAETHYLAMIN
 1979:  DIMETHYLFORMAMID
 2130:  DIMETHOXYETHAN
 2170:  TETRAHYDROFURAN
99997:  POLYPHOSPHORSAEURE
```

Diese lexikalische Auflistung der Stoffe enthält zuwenig Informationen. Eine für die Praxis wichtige Beziehung ist die eluotrope Reihe nach Trappe. (115) Sie ordnet die Substanzen nach steigender Elutionsfähigkeit an, als Adsorbens dient Aluminiumoxid. Jede vorläufig in dieser Reihe stehende Substanz erhält eine Kennzahl zwischen 1oo und 1.ooo. Das am wenigsten eluotionsfähige Lösungsmittel Petrolether wird durch die Zahl "1oo", die eluotionsfähigste Substanz Pyridin durch die Zahl "95o" gekennzeichnet.

Eluotrope Reihe nach Trappe

Substanz	Kennzahlvorschlag der Benutzer
Petrolether	1oo
Cyclohexan	15o
Schwefelkohlenstoff	2oo
Tetrachlormethan	25o
Dichlorethylen	3oo
Benzol	35o
Chloroform	4oo
Diethylether	45o
Tetrahydrofuran	5oo
Essigsäureethylester	55o
Aceton	6oo
Methylethylketon	65o
n-Butanol	7oo
Ethanol	75o
Methanol	8oo
Wasser	85o
Eisessig	9oo
Pyridin	95o
...	1ooo

Die Stellung eines Stoffes in dieser Reihe ist eindeutig festgelegt. Die verbleibenden Lücken können durch noch nicht in der Liste enthaltene Stoffe aufgefüllt werden. Da die Zuordnung von Kennzahlen als eine von mehreren Möglichkeiten zu verstehen ist, hat jeder Benutzer die Möglichkeit, sich seine besondere eluotrope Reihe zu definieren, die auf seine Bedürfnisse genau angepaßt ist.

Bei der Suche nach einem anderen Lösungsmittel kann sich der Anfrager in zwei Richtungen bewegen:

- die Reihe aufwärts, d.h. in Richtung der abnehmenden Polarität
  oder
- die Reihe abwärts, d.h. in Richtung der zunehmenden Polarität.

Die Abstände zwischen den Lösungsmitteln auf der numerischen Skala sind gleich, das heißt allerdings nicht, daß die Unterschiede in einer der Elutionskraft parallel laufenden physikalischen Eigenschaft in ebensogroßen Schritten abgestuft sind.
Neben den Eigenschaften, die sich in eine numerische Skala einordnen lassen, gibt es Eigenschaften, die man als Gegensatzpaare auffassen kann.

<u>Mögliche Gegensatzpaare bei der Beschreibung von Lösungsmitteln:</u>

- wassermischbar/nicht wassermischbar
- hygroskopisch/nicht hygroskopisch
- protisch/aprotisch
- polar/unpolar
- teuer/billig
- brennbar/nicht brennbar
- giftig/ungiftig
- deuteriert/nicht deuteriert
- technisch rein/pro analysi

Das Begriffspaar protisch/aprotisch kann dazu dienen, die Lösungsmittel zu klassifizieren. Im Vergleich zu einer numerischen Skala verliert man an Informationen, zusammen mit anderen Gegensatzpaaren erhält man brauchbare Beschreibungen.

Mit den Daten aus der eluotropen Reihe, dem Begriffspaar protisch/aprotisch und den Siedepunkten kann man z. B. folgende Fragen beantworten:

"Ich benötige ein Lösungsmittel, das mindestens genauso gut, wenn möglich besser löst als Diethylether und das einen höheren Siedepunkt besitzt."

"Welche hochsiedenden Lösungsmittel wurden schon einmal verwendet?"

Für die Weiterentwicklung des Systems kann auch die Aufnahme von Preisen für Lösungsmittel vorgesehen werden.

Vorschlag für ein Preisschema:

	DM/l
sehr teuer	5o
teuer	↑
preiswert	↓
billig	1

Abb. 21

Lösungsmittel für spektroskopische Zwecke bleiben bei diesen Überlegungen vorerst unberücksichtigt, da sie andere Anforderungen stellen.

4.4. Daten

4.4.1. Vorbemerkung

In den Naturwissenschaften fallen bei der experimentellen Arbeit große Mengen an Daten an: Elementarzusammensetzungen, Schmelzpunkte, Brechungsindices, Spektren, Beschreibungen von Reaktionen und von biologischen Eigenschaften. Daten werden unmittelbar als Ergebnisse von Experimenten produziert und dienen dazu, Fragestellungen zu lösen. Diese nicht unmittelbar zu verwendenden Daten dürften wohl die Hauptmenge der gemessenen Werte ausmachen: Um die Einheitlichkeit einer Substanzprobe zu überprüfen, werden eine Reihe von einander unabhängigen Methoden herangezogen, die die geforderte Sauberkeit der Probe für den beabsichtigten Zweck hinreichend genau belegen:

- Bestimmung des Schmelzpunkts oder des Mischschmelzpunkts,
- Messung des R_f-Werts bei einer chromatographischen Methode,
- Aufnahme eines Spektrums und Vergleich der Ergebnisse mit Vergleichsspektren aus der Datenbank.

Alle diese Daten werden nur gemessen, um eine Ja/Nein-Entscheidung zu treffen, die genauen numerischen Werte der Daten interessieren erst in zweiter Linie.

Diese erst mittelbar bedeutsamen Daten sind mit eine Ursache für das starke Ansteigen der publizierten Daten. Durch eine benutzergerechte Organisation der vorhandenen Datensammlungen kann man versuchen, die Daten besser zu nutzen.

Man sollte die Daten in zwei Gruppen einteilen:

- in nur für Spezialisten relevante Daten und
- in Daten, die für einen breiten Anwenderkreis wichtig sind.

Daten aus Naturwissenschaft und Technik können durch eine Einordnung in die folgenden Kategorien beschrieben werden:

- Meßbarkeit
  - wiederholt meßbar
  - einmalig meßbar
  - ortsabhängige Daten
  - ortsunabhängige Daten
- Unmittelbarkeit der Daten
  - experimentelle Daten
  - Daten, die sich aus Experimenten und aus Modellen ableiten lassen und
  - theoretische Vorhersagen
- Größe der Probe
  - individuelle Daten
  - statistische Daten

Weiter unterscheidet man Daten, die

- quantitativ
- semiquantitativ und
- qualitativ sind.

Sie präsentieren sich in Form von

- numerischen Werten,
- Graphen und Modellen und von
- Symbolen. (68)

Ein vorliegendes Datum kann gleichzeitig in mehrere dieser Facetten der Übersicht eingeordnet werden, z.B. der Schmelzpunkt einer Verbindungsprobe ist ein experimenteller Wert, der, unzersetztes Schmelzen der Probe vorausgesetzt, wiederholt gemessen werden kann, ortsunabhängig ist und eine quantitative Aussage ist, die sich in Form eines numerischen Werts präsentiert.

Die Ergebnisse einer Röntgenstrukturanalyse können als Zahlen vorliegen, die die Atomkoordinaten beschreiben oder als stereoskopisches Bild, das durch eine Analglyphenbrille betrachtet, einen dreidimensionalen Eindruck erweckt.

4.4.2. Überlegungen zur Datenstruktur in CHEMIS und die Erhebung der Daten

Betrachtet man die unterschiedlichen Anforderungen, die an die Daten von verschiedenen Seiten gestellt werden, so wird deutlich, daß das Problem vielschichtig ist:

- die EDV stellt strenge Anforderungen an die Formatierung der Daten: der erlaubte Wertebereich muß definiert sein, Transskriptionen bei den Bezeichnungen der Einheiten sind vorzunehmen und Steuerzeichen für die Tabellierung müssen definiert werden. Durch diese Festlegungen ist man sehr lange auf diese Konventionen fixiert, für trotzdem notwendige Änderungen ist der zu treibende Aufwand sehr groß.
- die Physik und die Meßtechnik verlangen den Nachweis von Größe mit Maßzahl und Einheit, Toleranzen und Meßfehlergrenzen sollten aus den Daten zu entnehmen sein.
- der Benutzer hat davon vollkommen andere Vorstellungen: er will die Daten leicht lesen können, zieht eine übersichtliche, am besten graphische Darstellung vor, möchte zuverlässige Daten aus dem System entnehmen können. Die Daten sollten so strukturiert sein, daß sie auf seine Fragen möglichst zutreffende Antworten geben.
  Diese Bedürfnisse erfordern eine benutzerfreundliche Aufarbeitung und Präsentation der Werte.

Es ist zwischen den Normen der Datenverarbeitung, den Erfordernissen der Physik und der Meßtechnik und den Benutzerbedürfnissen der organischen Chemiker ein tragfähiger Kompromiß zu finden.
Dieser Kompromiß wurde bisher so geschlossen, daß der Benutzer mit seinen Bedürfnissen an der letzten Stelle stand. Nicht nur, daß er sich mit unbequemen Datendarstellungsformen abfinden muß, er ist auch nicht in der Lage, die Daten so den Speichern zu entnehmen, wie er sie zu seiner Problemlösung benötigt. Er kann auch keine Manipulationen, wie z.B. die Umrechnung von Einheiten oder die Bildung von Mittelwerten vornehmen.

Man sollte deshalb versuchen, die Bedürfnisse der Benutzer stärker als bisher zu berücksichtigen, indem man

- die potentiellen Benutzer nach ihren Informationswünschen fragt,
- die so erfahrenen Bedürfnisse bei der Gestaltung des Systems mit einbezieht und
- die Programmierung des Systems so vornimmt, daß die Benutzerebene des Programms möglichst komfortabel ist. Die einzutippenden Befehle sollten sich möglichst wenig von der Fachsprache des Chemikers unterscheiden. Die Codes sollten mnemotechnisch günstig gebildet sein. Die Codes sollten so weit wie möglich selbsterklärend sein, so daß das notwendige Handbuch nur die ihm zugedachte Nebenrolle spielt.

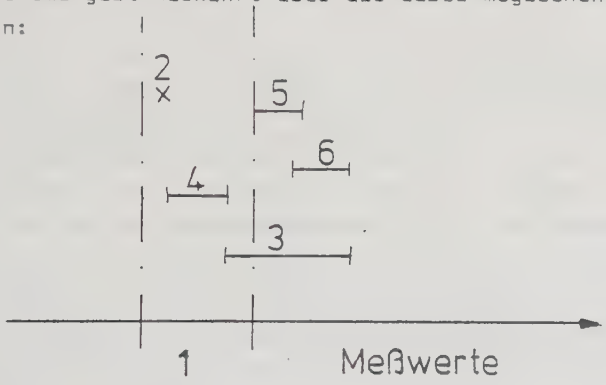
Ein Informationssystem, das zur Verarbeitung, Handhabung und Speicherung die Leistungsfähigkeit eines Computers benutzt, wird auf Randbedingungen der EDV Rücksicht nehmen müssen.

Die Daten wurden aus dem Bereich der Chlorformamidin-Chemie ausgewählt und ohne Bewertung eingespeichert. Sie wurden aus der zugänglichen Literatur entnommen oder durch direkte Befragung der Chemiker gewonnen. Es wurden Datenerfassungsschemata entworfen, die dazu dienen, die Daten aus den verschiedenen Darstellungsformen in der Literatur oder in den Laborjournalen in einheitliches Schema zu überführen, das den Erfordernissen der EDV angepaßt ist.
Diese Datenerfassungsschemata sind nicht starr zu sehen, sondern als Versuch aufzufassen, die Datenerhebung vorläufig zu strukturieren und vor allem die verschiedenen Datenarten und -kategorien vergleichbar zu machen.

Die Bewertung von nachgewiesenen Daten ist kein Ziel dieser Systementwicklung. Um diese Aufgabe zu bewältigen, müßte eine Menge zusätzlicher Literaturarbeit geleistet werden, es müßten Vergleichsdaten beschafft werden, die Daten müßten zusammengestellt und miteinander verglichen werden. Bei den teilweise neuen Stoffen, die im Laufe der Zeit eingespeichert wurden, fehlen die Vergleichsdaten vollständig.

Es liegt im wohlverstandenen Eigeninteresse des Systembenutzers, dafür zu sorgen, daß nur zuverlässige Daten in das System aufgenommen werden und daß zweifelhafte Daten bis zur Klärung des Sachverhalts ausgeschlossen bleiben. Der Benutzer selbst hat also die Zuverlässigkeit seines Datenbestandes in der Hand.

Das System CHEMIS erlaubt die Angabe eines Meßwertbereichs sowohl bei der Eingabe der Daten als auch bei der Formulierung der Abfrage.
Das Schema gibt Auskunft über die dabei möglichen Konstellationen:



1 durch die Formulierung der Abfrage vorgegebener Wertebereich
2 konkreter Meßwert
3 Meßwertbereich innerhalb und außerhalb der Toleranzzone
4 Meßwertbereich innerhalb der Toleranzzone
5 Meßwertbereich an der Grenze der Zone
6 Meßwertbereich außerhalb des Bereichs

Unstrittig sind die Fälle 2 und 4, bei Fall 3 treten ebenfalls keine Schwierigkeiten auf.
Fall 6 fällt außerhalb des angefragten Bereichs und wird in CHEMIS nicht nachgewiesen.
Schwierigkeiten können im Fall 5 auftreten, da nur die Grenze der Bereiche übereinstimmen. Für einen Nachweis könnte sich ein Problem ergeben: Ist es gerechtfertigt, diesen Wertebereich mit einzubeziehen oder muß man ihn wegen zu geringer Übereinstimmung verwerfen?

Im System CHEMIS ist die Abfrage so programmiert, daß auch der Fall 5 als zutreffend nachgewiesen wird.
Der Ballastanfall dürfte sich in Grenzen halten, die Sicherheit, ein vollständiges Ergebnis zu erhalten, steigt an.

Fragebereich	Wertebereich	
	nachgewiesen	nicht nachgewiesen
1	2 3 4 5	6

Sollte trotzdem bei einer Fragestellung keine Antwort zu erhalten sein, kann es daran liegen, daß
- die Verbindung noch nicht im System eingespeichert ist

oder daß
- die Verbindung zwar vorhanden ist, aber der Datensatz fehlt.

Als dritte Möglichkeit bleibt noch, daß zwar Verbindung oder Reaktion mit allen Daten vorhanden ist, daß aber die Bereiche der Abfrage so gewählt sind, daß keine Übereinstimmung erzielt wird.
Um diesem Nachteil begegnen zu können, sind die Bereichsgrenzen bei der Abfrage variabel gehalten. Der Benutzer kann also entscheiden, ob er lieber genau fragen will und einen gewissen Informationsverlust in Kauf nehmen möchte oder ob er eine mehr oder weniger breite Frageformulierung zuläßt, die ihm einen höheren Ballastanteil liefert.
Vom System fest vorgegebene Bereichsgrenzen sind in diesem Anwendungsfall nicht sinnvoll.

4.4.2.1. Daten zu Verbindungen

Jede in CHEMIS eingespeicherte Verbindung wird mit einer fortlaufenden Nummer gekennzeichnet. Diese Nummer hat keinerlei strukturelle Bedeutung. Sie dient lediglich als Ordnungsmerkmal und als Verknüpfungshilfsmittel für die verschiedenen Dateien, die einer Substanz zugeordnet werden können.
Der Nummernkreis von 001 bis 1 099 ist für anorganische Stoffe vorgesehen, ab 1 100 beginnt der Bereich der organischen Stoffe. Gemische und in ihrer Zusammensetzung nicht genau definierte Substanzen belegen den Bereich von 99 998 an, wobei die Zählung rückwärts läuft.
Die Namen der chemischen Verbindungen sind das nächste Beschreibungselement. Die Namen werden aus den Veröffentlichungen übernommen, offensichtliche Fehler oder Ausdrücke aus dem Laborjargon werden korrigiert.
Die Summenformel eines Stoffes wird nach dem System von Hill aufgelistet. (116)
Der Schmelzpunkt einer Verbindung kann auf 0,1 ^{o}C genau referiert werden. Ebenso ist die Nennung eines Schmelzbereichs erlaubt. Er kann beliebig groß sein.
Es kann notwendig werden, die Zahlenangaben zum Schmelzpunkt durch Kommentare zu erläutern.
Zur Zeit stehen zur Verfügung:

- z : Zersetzung und
- ö : für Öl.

Die Abkürzungen werden nur bei der Eingabe verwendet, im Ausdruck der Listen oder auf dem Bildschirm erscheint der Klartext.
Der Siedepunkt eines Stoffes wird analog behandelt.
Der Druck, bei dem der Siedepunkt gemessen wird, kann in Atmosphären oder in Torr wiedergegeben werden, für die Abfrage stehen beide Einheiten zur Verfügung, intern wird umgerechnet.

Die Werte für die Basen- und Säurekonstanten werden aus der Literatur übernommen.
Da die Beziehung

$$pK_b + pK_s = 14$$

gilt, kann man die Werte ineinander umrechnen.
Um die notwendige Eindeutigkeit bei den Bezugsgrößen zu erreichen, sollte man sich auf eine Konstante einigen.

Aus praktischen Gründen wurde die Basenkonstante als Bezugsgröße gewählt. Je stärker also eine Base ist, desto größer ist der Zahlenwert der Basenkonstante.

Die Angabe einer Familienkennziffer hat den Zweck, gleichartige Stoffe zusammenzufassen. In dieser Experimentalphase sind folgende Familien definiert:

F 1oo: Peptide
F 2oo: Säuren
F 3oo: Halogenide
F 4oo: Carbonsäureester
F 5oo: Olefine
F 6oo: Harnstoffe
F 61o: Chlorformamidine
F 62o: Isocyaniddichloride

Die Familienkennziffer ermöglicht eine einfache und unkomplizierte Abfrage nach wichtigen Verbindungsklassen, ohne das reaktive Verhalten der Verbindungen im einzelnen im System abgespeichert haben zu müssen. (Reaktionsgleichungen)
Dieses Hilfsmittel zur Zusammenführung chemisch zusammengehöriger Stoffe sollte im Lauf der Zeit immer mehr überflüssig werden, da die Beziehungen zwischen den Stoffen durch Reaktionsgleichungen beschrieben und nicht durch eine - letztendlich doch willkürliche Definition der Familien - festgelegt werden.

Vorschlag für ein Erfassungsschema für Daten chemischer Verbindungen:

exec chvein				
Nr.:				
Name:				
Summenformel:				
Fp.:		°C		
Bem. Fp.:				
Kp.:		°C		
Bem. Kp.:				
Druck Kp.:		atm		Torr
Familie:				
Basenkonstante:				

Beispiel für eine Verbindung ohne weitere Daten:

```
ei,v 1
     1:  CHLOR
         Cl 2
```

Beispiel für eine Verbindung mit ausführlichem Datensatz:

```
ei,v 1767
 1767:  PHENYLISOCYANIDDICHLORID
        C 7, H 5, N, Cl 2
        FAMILIE:          620 ISOCYANIDDICHLORID
        SIEDEPUNKT:       103 - 106 GRAD C  BEI 30 TORR
        I R - S P E K T R U M   (ALDRICH-KATALOG 1975):
        MEDIUM:         KALIUMBROMID
        WELLENZAHL    INTENSITAET    ZUORDNUNG
          3050          SCHWACH      C-H-VALENZSCHWINGUNG
                                     AROMAT
          1850          SEHR STARK   C-N-DOPPELBINDUNGSVALENZSCHWINGUNG
                                     ISOCYANIDDICHLORID
           920          STARK        --
           880          STARK        --
           755          STARK        C-H-WAGGINGSCHWINGUNG
                                     5 H
           690          STARK        C-H-WAGGINGSCHWINGUNG
                                     5 H
```

Beispiel für eine Verbindungsfamilie:
Teilliste der Harnstoffe

```
?
f 600
LISTE DER HARNSTOFFE:
 2168:  3-CHLOR-N-DIMETHYLCARBAMOYL-6-METHYLBENZO(b)THIOPHEN-2-CARBOXAMI
D
 2167:  3-CHLOR-6-METHYLBENZO(b)THIOPHEN-2-CARBOXAMID
 2057:  N-(3-CHLOR-6-METHYL-2-CARBONYLBENZO-b-THIOPHENO)-N´-DIMETHYLHARN
STOFF
 1863:  N-(4-CHLORPHENYL)-N´-DIMETHYLHARNSTOFF
 2178:  3-CHLOR-N-(PIPERIDINOCARBONYL)BENZO(b)THIOPHEN-2-CARBOXAMID
 1816:  N.N´-DI-(3-CHLORPHENYL)-THIOHARNSTOFF
 1612:  N.N´-DICYCLOHEXYLHARNSTOFF
 1807:  (4-DIETHYLAMINO-PHENYL)-CARBAMIDSAEUREMORPHOLID
 1835:  N.N´-DIISOPROPYLHARNSTOFF
 1837:  N.N´-DIISOPROPYLTHIOHARNSTOFF
 1803:  N.N´-(3.3´-DIMETHOXYBIPHENYL-4.4´-DIYL)-bis-CARBAMIDSAEURE-DIMOR
PHOLID
 1819:  N.N´-DI-(4-METHOXYPHENYL)-THIOHARNSTOFF
 1814:  N.N´-DIMETHYLHARNSTOFF
 1858:  N-DIMETHYL-N´-METHYLTHIOHARNSTOFF
 1834:  N.N´-DIMETHYLTHIOHARNSTOFF
 1854:  N.N´-DI-(beta-NAPHTHYL)-THIOHARNSTOFF
 1838:  N.N´-DIOCTADECYL-HARNSTOFF
 1973:  N.N´-DIPHENYLHARNSTOFF
 1840:  N.N´-DIPHENYL-N´-ISOPROPYL-THIOHARNSTOFF
 1844:  N.N´-DIPHENYLTHIOHARNSTOFF
 1869:  DI-PYRROLIDINHARNSTOFF
 1934:  DODEKADEUTEROTETRAMETHYLHARNSTOFF
 1802:  (4-ETHOXY-1-NAPHTHYL)-CARBAMIDSAEUREMORPHOLID
 2061:  N-(6-METHYL-3-CHLOR-2-CARBONYLBENZO-b-THIOPHENO)-N´-DIMETHYLTHIO
HARNST
        OFF
 1801:  N-(4-METHYLMERCAPTOPHENYL)-CARBAMIDSAEUREMORPHOLID
 1998:  N-METHYL-N-PHENYL-(3)-PYRIDYLHARNSTOFF
 2157:  3-(N´-METHYL-N´-PHENYL-UREIDO)-THIOPHEN-2-CARBONSAEURE-METHYLEST
ER
 2011:  4-METHYLPIPERAZIN-1-CARBONSAEUREFURFURYLAMID
 2159:  3-(4-METHYL-PIPERAZIN-1-CARBONYLAMINO)-THIOPHEN-2-CARBONSAEURE-M
ETHYLE
        STER
 2074:  1.5-bis-(MORPHOLIN-4-CARBAMIDO)-NAPHTHALIN
 1918:  MORPHOLIN-4-CARBONSAEURE-(4-CHLOR-PHENYL)-AMID
 1917:  MORPHOLIN-4-CARBONSAEURE-(3.5-DIBROM-4-BIPHENYL)-AMID
 1794:  MORPHOLIN-4-CARBONSAEURE-1.2-DIPHENYL-VINYLAMID
 2008:  MORPHOLIN-4-CARBONSAEURE-FURFURYLAMID
 2005:  MORPHOLIN-4-CARBONSAEURE-2-FURYLAMID
 1922:  MORPHOLIN-4-CARBONSAEURE-(3-NITROPHENYL)-AMID
 1991:  MORPHOLIN-4-CARBONSAEURE-(3)-PYRIDYLAMID
 1786:  MORPHOLIN-4-CARBONSAEURE-trans-STYRYLAMID
 2021:  MORPHOLIN-4-CARBONSAEURE-(2)-THIENYLAMID
 2156:  3-(MORPHOLIN-4-CARBONYLAMINO)-THIOPHEN-2-CARBONSAEURE-METHYLESTE
```

...

4.4.2.2. Infrarot-Spektren

Die Infrarotspektroskopie ist eine aus dem organischen Laboratorium nicht mehr hinwegzudenkende Untersuchungsmethode. Die Geräte sind sehr weit verbreitet, es liegen umfangreiche Sammlungen von Spektren vor, die als Vergleichsmaterial dienen können. (117 - 119, 120)
Ihre weite Verbreitung hat folgende Gründe:

- Die Geräte sind relativ preiswert.
- Die Präparation der Proben erfordert wenig Aufwand.
- Es sind wenige und dann preiswerte Hilfsstoffe erforderlich.
- Die zu messende Substanz kann gasförmig, flüssig oder fest sein, Lösungen sind ebenfalls meßbar.
- Der Zeitaufwand zur Registrierung des Spektrums ist kurz.
- Die Methode ist empfindlich genug, um auch kleine Mengen zu messen. (ca. 1 mg)
- Die Spektren sind auch für den Anfänger und Nichtfachmann interpretierbar. Tabellen helfen bei dieser Aufgabe.

Dem Denken der Chemiker in funktionellen Gruppen kommt die Eigenschaft der Moleküle entgegen, charakteristische Gruppenschwingungen im Spektrum zu zeigen. Gedanklich zerlegt man das Molekül in mehrere voneinander unabhängig schwingende Teile, die sich - in erster Näherung - nicht beeinflussen. Da die Energieunterschiede zwischen den einzelnen Schwingungselementen genügend groß sind, ist dieses Verfahren erlaubt. Das Gesamtspektrum eines Moleküls erhält man dann durch Kombination der Einzelschwingungen aller im Molekül vorhandenen funktionellen Gruppen.

Vorschlag für ein Erfassungsschema für IR-Daten:

exec chirein				
Nr.:		Gerät:		Medium:
	Lage	Intensität	Element	Interpretation
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				

Die Bandenlage wird zwar durch die strukturellen Besonderheiten des Moleküls modifiziert, aber nicht so sehr, daß keine Regelmäßigkeiten meht zu erkennen wären. Diese Feinverschiebungen dienen zur genaueren Analyse des Moleküls. Die funktionellen Gruppen in der Struktur eines Moleküls finden also ihre Entsprechung in der Bandenstruktur des Infrarotspektrums. (121)

Wieviele einzelne Interpretationsmerkmale und Beschreibungselemente sind bei der Konzeption zu berücksichtigen ?

Eine Zählung der Interpretationsmerkmale in einem Spektroskopie-Buch ergibt ca. 2oo verschiedene Zuordnungen unterschiedlicher Genauigkeit und Spezifität.

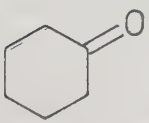
Die Menge an Zuordnungen ist zu groß, um in einer einzigen Liste aufgeführt werden zu können. Eine Liste sollte nicht wesentlich länger als zwei Bildschirmseiten sein, damit die Übersichtlichkeit erhalten bleibt. Eine Unterteilung der Interpretationen in eine Beschreibung der schwingenden Bindung und in die modifizierende Umgebung liegt nahe und befindet sich auch in Übereinstimmung mit der oben skizzierten heuristischen Denkweise bei der Interpretation der Spektren.

Liste der Interpretationen für IR-Spektren	
Liste der Bindungen	Liste der Schwingungselemente

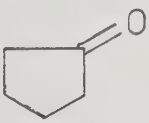
Die Gesamtliste zerfällt in zwei Bereiche unterschiedlicher Spezifität. Eine Liste gibt über die auftretenden Schwingungselemente Auskunft, der Informationsgehalt eines IR-Spektrums ist aber mit diesen Angaben noch nicht adäquat ausgeschöpft. Die Angaben müssen noch modifiziert werden.

Das geschieht in der Liste der einzelnen Schwingungselemente.

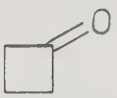
Die strukturellen Einflüsse haben zur Folge, daß man bei der Kohlenstoff-Sauerstoff-Doppelbindungsschwingung zwischen folgenden Strukturen genau unterscheiden kann:



1705 - 1725 cm^{-1}



1740 - 1750 cm^{-1}



ca. 1775 cm^{-1}

Diese Atomgruppen führen alle zu starken Banden, die im Spektrum sofort auffallen und daher einfach zu interpretieren sind.
Im Bereich zwischen ca. 1 400 und ca. 900 cm^{-1} erscheinen im Spektrum eine Vielzahl von Banden, die nicht interpretierbar sind, aber doch für eine bestimmte Substanz charakteristisch sind, da nur unbedeutende Änderungen in der Struktur deutlich erkennbare Veränderungen in diesem sog. "fingerprint"-Bereich zur Folge haben. Ein Spektrennachweissystem kann sich also nicht nur auf den Nachweis von interpretierten Banden beschränken, sondern es muß auch die Möglichkeit offen lassen, Banden ohne Angabe ihrer Interpretation aufzulisten. Dieser Tatsache wird durch die Möglichkeit Rechnung getragen, die Angaben über Schwingungselement und Interpretationsmerkmal auch weglassen zu können.
Die Intensität wird nur grob quantitativ nachgewiesen:

- schwache Banden,
- mittlere,
- starke und
- sehr starke Banden werden unterschieden.

Da die meisten IR-Spektren qualitative Spektren sind, die zur Messung gelangende Probe besteht aus einem Preßling aus einem innigen Gemisch aus etwa 1 - 1,5 mg Substanz und ca. 300 mg KBr, wird auf eine genauere Intensitätsangabe verzichtet.

Da die Nullinie nicht korrigiert ist und der Chemiker nur an den qualitativen Daten interessiert ist, reicht dieses grobe Raster aus. Die Bandenform, ob scharf, breit oder mit einer Schulter, wird ebenfalls nicht nachgewiesen, da diese Angaben oft mit einer gewissen Willkürlichkeit behaftet sind.

Eine IR-Bande in CHEMIS besteht aus folgenden Datenelementen:
- der Wellenzahl
- der Intensität
- dem Schwingungselement und
- der Interpretation.

Um eine Substanz ausreichend zu charakterisieren, werden zehn Banden als ausreichend angesehen. Man kann sicher wesentlich mehr Banden einem Spektrum entnehmen, interpretieren wird man schon weniger, zumal der Aufwand zur Interpretation der kleineren Banden sehr hoch ist.
In den Publikationen wird häufig nur eine einzige, besonders charakteristische Bande aufgeführt, die Intensitätsangabe fehlt dann ebenfalls.

Anzahl der Banden	Quelle
1oo - 2oo	Originalspektrum
3 - 5	interpretiertes Spektrum in der Originalarbeit
1	Literatur

Das Informationssystem CHEMIS muß sich zwischen den Extremen des vollständigen Nachweises der Banden und nur einer einzigen aufgeführten Bande bewegen und erreichen, daß die getroffene Auswahl die ausreichende Menge an Informationen vermittelt.

In einem Vorspann zu jeden Spektrum wird

- über die untersuchte Verbindung Auskunft gegeben. Verbindungen werden durch die Verbirdungsnummer bei der Eingabe verschlüsselt. Es können nur Spektren von Stoffen nachgewiesen werden werden, die dem System schon bekannt sind. Spektren neuer Stoffe werden durch eine Abfrage mit dem Datenbestand verglichen.
- das Medium oder das Einbettungsmittel beschrieben. Es hat einen geringen Einfluß auf das Spektrum, ob eine Substanz in einer Nujol-Verreibung oder einem KBr-Preßling gemessen wird. Banden, die auf polare Gruppen zurückzuführen sind, werden verschoben.
- Die als Matrix verwendete Verbindung wird ebenfalls über ihre Verbindungsnummer verschlüsselt, eine eigne Liste mit allen Einbettungsmitteln wird nicht definiert, der Aufwand hierfür erscheint zu groß.
- Das Meßgerät oder die sonstige Quelle, aus der das Spektrum entnommen wurde, wird durch Schlüsselzahlen nachgewiesen.

Liste der Schlüsselzahlen	
1	Perkin-Elmer 157
2	Perkin-Elmer 117
3	Aldrich-Katalog, (1977)
4	Quelle unbekannt
5	Verweis auf die Literaturstelle
6	...

Wenn aus der Quelle das Meßgerät nicht zu entnehmen ist, so wird das Feld "Meßgerät" mit der Schlüsselzahl "4" ausgefüllt.
Eine strenge Trennung des Inhalts im Datenfeld "Meßgerät" in eine Liste der Meßgeräte und in eine Literaturliste durch die Definition zweier Unterdateien

- "Meßgeräte"
  und
- "Quellen"

erscheint nicht sinnvoll, da es für die praktischen Bedürfnisse des Laborchemikers unerheblich ist, zu wissen aus welchen genauen Quellen seine Daten stammen, solange er den Weg zum Ursprung der Daten einfach nachvollziehen kann.
Neben der benutzergerechten Aufarbeitung und Präsentation der Informationen sollte bei der Systementwicklung darauf geachtet werden, daß der laufende Betrieb des Systems mit so wenig wie möglich Aufwand zu gewährleisten ist.
Der Benutzer ist nur willens, die unbedingt notwendigen Arbeiten auszuführen. Die Pflege des Systems sollte sich möglichst wenig von der gewohnten Führung des Laborjournals unterscheiden.
Daher ist die Anzahl der Listen, die

- zu eröffnen sind und deren Inhalt zu definieren ist,
- laufend zu ergänzen und
- abzufragen sind,

möglichst gering zu halten.

Es ist nun noch nachzuprüfen, ob der so verschlüsselt dargestellte Inhalt der IR-Spektren den Informationsbedarf der Chemiker erfüllen kann.
Vorab ist zu klären, ob mit der vorhandenen Zahl der Verschlüsselungselemente eine ausreichende Zahl von Stoffen beschreibbar ist, so daß eine umkehrbar-eindeutige Zuordnung von verschlüsselt dargestelltem Spektrum und Stoff existiert. Im umfangreichsten Fall benötigt man einen Schlüssel, um $6 \cdot 10^6$ Verbindungen zu unterscheiden.
Man benötigt 23 verschiedene Merkmale, um die bisher bekannten Stoffe zu codieren.
Die Zahl der in CHEMIS zur Verfügung stehenden Beschreibungselemente übersteigt diese Zahl.
Man hat zur Verfügung:

- 23 Schwingungselemente
- 80 Interpretationsmerkmale (gemeinsam für IR-Spektren und für NMR-Spektren)
- 4 Intensitäten und
- Wellenzahlen von 4 000 cm^{-1} bis 400 cm^{-1}, 3 600 unterscheidbare Merkmale. (Auflösung: 1 cm^{-1})

Das Problem der mangelnden Unterscheidbarkeit tritt also nicht auf, man hat sogar eine redundante Darstellungsform gewählt, so daß man die Möglichkeit hat, einzelne Elemente wegzulassen, wenn sie sich als nicht wesentlich erweisen, ohne in Gefahr zu laufen, daß die Unterscheidbarkeit der Verbindungen verloren geht.

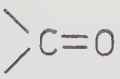
Das umseitig abgebildete Erfassungsschem für infrarotspektroskopische Daten hat zwei Aufgaben:

- es soll das Extrahieren der Daten aus der Literatur erleichtern und
- es soll als Entwurfsmuster für die Gestaltung der Masken auf dem Bildschirm dienen.

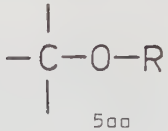
Liste der Schwingungselemente:

se	
1	C-C-EINFACHBINDUNGSVALENZSCHWINGUNG
10	C-C-DOPPELBINDUNGSVALENZSCHWINGUNG
20	C-C-DREIFACHBINDUNGSVALENZSCHWINGUNG
30	C-O-EINFACHBINDUNGSVALENZSCHWINGUNG
40	C-O-DOPPELBINDUNGSVALENZSCHWINGUNG
50	C-N-EINFACHBINDUNGSVALENZSCHWINGUNG
60	C-N-DOPPELBINDUNGSVALENZSCHWINGUNG
70	C-N-DREIFACHBINDUNGSVALENZSCHWINGUNG
80	C-H-VALENZSCHWINGUNG
90	C-H-DEFORMATIONSSCHWINGUNG
100	C-H-WAGGINGSCHWINGUNG
110	O-H-VALENZSCHWINGUNG
120	N-H-VALENZSCHWINGUNG
130	S-H-VALENZSCHWINGUNG
140	C-S-VALENZSCHWINGUNG
150	C-Si-VALENZSCHWINGUNG
160	C-Halogen-VALENZSCHWINGUNG
170	C-Nitro-VALENZSCHWINGUNG
180	C-P-VALENZSCHWINGUNG
190	C-C-GERUESTSCHWINGUNG
200	N-N-VALENZSCHWINGUNG
210	P-Aryl-VALENZSCHWINGUNG
220	P-O-DOPPELBINDUNGSVALENZSCHWINGUNG

Beispiele:

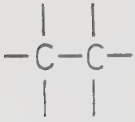


23o

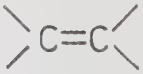


5oo

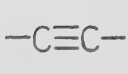
oder:



1oo



13o



14o

Liste der Interpretationsmerkmale:

int	
100	ALIPHAT
110	AROMAT
111	5 H
112	4 H
113	3 H
120	ALKAN
121	METHYLGRUPPE
122	METHYLENGRUPPE
123	METHOXYGRUPPE
130	ALKEN
140	ALKIN
200	CARBONYLVERBINDUNGEN
210	ALDEHYD
211	ALKYLALDEHYD
212	ARYLALDEHYD
230	KETON
231	alpha.beta-ungesaettigtes KETON
232	ARYLKETON
233	alpha-HALOGENKETON
234	alpha-DIKETON
235	diungesaettigtes KETON
236	DIHALOGENKETON
237	4-RINGKETON
238	5-RINGKETON
239	6-RINGKETON
240	THIOKETON
260	CHINON
300	ALKOHOL
301	primaerer ALKOHOL
302	sekundaerer ALKOHOL
303	tertiaerer ALKOHOL
310	PHENOL
400	AMIN
401	primaeres AMIN
402	sekundaeres AMIN
403	tertiaeres AMIN
404	AMINSALZ
405	HYDRAZINDERIVAT
410	aromatisches AMIN
420	AMID
421	LACTAM

Liste der Interpretationsmerkmale:

(Fortsetzung)

430 AMINOSAEURE
431 AMINOSAEUREESTER
432 AMINOSAEUREHYDROCHLORID
441 PEPTIDE
442 DIPEPTIDE
443 TRIPEPTIDE
450 NITRIL
451 ISOCYANAT
452 ALKYLNITRIL
453 aromatisches NITRIL
454 ISOCYANIDDICHLORID
460 AZID
470 HARNSTOFF
471 THIOHARNSTOFF
472 ISOHARNSTOFF
473 CHLORFORMAMIDIN
480 AMIDIN
500 ETHER
501 aliphatischer ETHER
502 aromatischer ETHER
503 cyclischer ETHER
600 CARBONSAEURE
601 CARBONSAEUREHALOGENID
602 CARBONSAEUREANHYDRID
603 alpha-HALOGENSAEURE
620 ESTER
621 LACTON
700 HETEROCYCLUS
701 PYRIDIN
702 PYRIMIDIN
703 MORPHOLIN
704 THIOPHEN
800 HALOGENIDE
801 FLUORID
802 CHLORID
803 BROMID
804 IODID
820 TRIPHENYLPHOSPHAN
821 TRIPHENYLPHOSPHANOXID

Beispiel für ein Spektrum eines Einbettungsmittels:

```
ir,v 99998
99998:  NUJOL
        I R - S P E K T R U M   (ALDRICH-KATALOG 1975):
        MEDIUM:          NATRIUMCHLORID
        WELLENZAHL    INTENSITAET    ZUORDNUNG
          2950          SEHR STARK    --
          2850          STARK         --
          1450          MITTEL        --
          1380          SCHWACH       --
           720          SCHWACH       --
```

Beispiele für IR-Spektren:

Spektrum mit nur einer Bande:

```
ir,v 1956
 1956:  SCHWEFELKOHLENSTOFF
        I R - S P E K T R U M   (PERKIN-ELMER 177):
        MEDIUM:          KALIUMBROMID
        WELLENZAHL     INTENSITAET     ZUORDNUNG
          1500          SEHR STARK     C-S-VALENZSCHWINGUNG
                                       THIOKETON
```

Spektrum mit der maximalen Anzahl der Banden:

```
ir,v 1785
 1785:  TRIPHENYLPHOSPHANOXID
        I R - S P E K T R U M   (PERKIN-ELMER 177):
        MEDIUM:          KALIUMBROMID
        WELLENZAHL     INTENSITAET     ZUORDNUNG
          3060          SCHWACH        C-H-VALENZSCHWINGUNG
                                       AROMAT
          1590          SCHWACH        C-C-DOPPELBINDUNGSVALENZSCHWINGUNG
                                       AROMAT
          1485          MITTEL         C-C-DOPPELBINDUNGSVALENZSCHWINGUNG
                                       AROMAT
          1460          SEHR STARK     --
          1190          SEHR STARK     --
          1120          SEHR STARK     --
           995          MITTEL         --
           750          MITTEL         --
           720          SEHR STARK     C-H-WAGGINGSCHWINGUNG
                                       5 H
           695          SEHR STARK     C-H-WAGGINGSCHWINGUNG
                                       5 H
```

4.4.2.3. Kernresonanzspektren

Von den prinzipiell zur Verfügung stehenden Kernen, die ein paramagnetisches Moment haben, sind die Kerne von Wasserstoffatomen am geeignetsten, um Aussagen über die Struktur organischer Verbindungen zu liefern. Die Meßmethode ist empfindlich, einfach und schnell. Daher ist die Kernresonazspektroskopie neben der IR-Spektroskopie mit die am weitetsten verbreitete Methode zur Strukturaufklärung.
Bedingt durch

- günstiges Signal/Rausch-Verhältnis,
- einfache Probenvorbereitung,
- kurze Meßzeiten und
- quantitative Meßbarkeit der relativen Protonenzahlen in einem Molekül,

sind Aussagen möglich, die die Ergebnisse der IR-Spektroskopie ergänzen.
Unterstrichen wird die besondere Eignung der Wasserstoffkerne als Objekte des Kernresonanzexperiments durch die hohe relative Empfindlichkeit im Vergleich zu anderen Atomkernen:

Kernsorte	relative Empfindlichkeit
^{1}H	1
^{2}D	$1,5 \cdot 10^{-6}$
^{13}C	$1,7 \cdot 10^{-4}$
^{14}N	10^{-3}
^{15}N	10^{-5}
^{31}P	$6,7 \cdot 10^{-2}$

(Magnetfeld: 2,3 Tesla, Protonenresonazfrequenz: 100 MHz)

Vorschlag für ein Erfassungsschema von NMR-Daten:

exec chnmrein					
Verbindungsnummer:					
Nr. des Spektrums:					
Messfrequenz:		MHz			
Standard:					
Temperatur:		°C			
Lösungsmittel:					
Peak	δ [ppm]	H	Auf-spaltung	Hz	Zu-ordnung
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					

IR-Spektroskopie und NMR-Spektroskopie liefern einander ergänzende Aussagen. Die Infrarotspektren liefern genaue Informationen über die funktionelle Gruppen, die zu Absorptionen Anlaß geben. Sie liefern auch Gerüstschwingungsabsorptionen, diese sind aber nicht so einfach zu interpretieren wie die C-O-Doppelbindungsvalenzschwingung im Bereich um 1 7oo cm^{-1}.

Einige Beispiele sollen die unterschiedlichen Heuristiken erläutern, die den Interpretationen zu Grunde liegen:

- Carbonylverbindungen: Die Absorptionsfrequenz der Carbonylschwingung liefert einen Hinweis auf die Ringgröße. Die näheren Informationen über die Umgebung erhält man aus der NMR-Spektrum der Substanz, man kann die Protonen in der Nachbarschaft der Carbonylgruppe zählen und über die Messung der Kopplungskonstante die Stereochemie aufklären.
- Alkohole: Die Absorption der O-H-Schwingung tritt im IR-Spektrum oftmals sehr breit und daher uncharakteristisch auf, so daß die Zuordnung zu der Hydroxylgruppe nicht immer eindeutig möglich ist, da die Restfeuchte in der Matrix zu der gleichen Absorption Anlaß gibt. In der NMR-Spektroskopie kann man das Problem durch Austausch des Hydroxylprotons mit einem Überschuß an D_2O lösen: die der Hydroxylgruppe zuzuordnenden Signale verschwinden, durch Integration und Vergleich mit dem Ursprungsspektrum erhält man die Anzahl der Hydroxylprotonen.
- Amine: Das Vorgehen bei der Untersuchung von Aminen unterscheidet sich nicht wesentlich von den Methoden bei der Untersuchung von Alkoholen. Da aber die Einstellung des Austauschgleichgewichts behindert sein kann, ist es notwendig, die Einstellung des Gleichgewichts durch eine Spur Trifluoressigsäure zu katalysieren.

Jedes Kohlenstoffatom, das entweder

- mit Wasserstoff substituiert ist oder
- mit einer funktionellen Gruppe verbunden ist, kann mit mindestens einer der beiden Methoden untersucht werden.

Ohne den Umweg über die Bindungspartner der Kohlenstoffatome ist der Weg mit der ^{13}C-Kernresonanzspektroskopie. In dieser Arbeit wird diese Methode nicht näher benandelt, da die Ergebnisse des Meßverfahrens einer Darstellung der Struktur bedürfen, die das System CHEMIS aber zur Zeit noch nicht bieten kann, da es noch nicht über einen graphischen Bildschirm verfügt.

Darstellung der Meßdaten:

Im Gegensatz zur IR-Spektroskopie sind die Ergebnisse der NMR-Spektroskopie weniger geräteabhängig, solange man sich in einer Leistungsklasse der Geräte bewegt. Die Auflösung der Geräte hängt von dem Magnetfeld ab: Je stärker das Magnetfeld, desto besser die Auflösung der Signale. Die Magnetfeldstärke bestimmt die Meßfrequenz.
Zur Charakterisierung dient üblicherweise die Größe der Meßfrequenz:

- 60 MHz - Geräte für den Routinebetrieb
- 90 MHz
- 100 MHz
- 270 MHz - supraleitende Magnete, FT-Messung

(Frequenzen für Protonen)

Wie bei den Infrarot-Spektroskopie dient die Verbindungsnummer als Einstieg und zugleich als Bindeglied, über das die einer Verbindung zugeordneten Daten miteinander verknüpfbar sind. Die Verbindungsnummer erfüllt die Funktion eines Gelenks zwischen den einzelnen Dateien, daher kann man ohne großen Aufwand sich neue Dateien und Listen definieren, falls es sich als notwendig erweisen sollte.

Pro Verbindung können zwei Spektren abgespeichert werden. Nur so können Spektren erfaßt werden, die durch H/D-Austausch erhalten werden oder die bei unterschiedlichen Temperaturen gemessen werden. Die andere Möglichkeit, durch eine Modifikation an dem jeweils betroffenen Peak eine Zusatzinformation zu verschlüsseln, wurde nicht gewählt, da die Spektrendarstellung zu unübersichtlich werden würde. Das erste Spektrum ist immer das Spektrum ohne Modifikationen, im zweiten Spektrum hat dann der Austausch oder die Temperaturänderung stattgefunden. Es ist nicht ausreichend, nur die Peaks aufzuführen, die von der Änderung der Meßbedingungen unmittelbar betroffen sind. Durch Änderung des Lösungsmittels ergeben sich Verschiebungen der Peaklagen bis zu o,2 ppm. Die Lage der neu entstehenden HDO-Peaks wird nicht nachgewiesen.
Die Meßfrequenz wird in der Einheit "MHz" aufgeführt.
Im Datenfeld "Standard" wird die Verbindung mit ihrer Verbindungsnummer verschlüsselt, die als (interner oder externer) Standard dient. In den meisten Fällen in der organischen Chemie wird dies Tetramethylsilan sein.
Die Temperatur, bei der die Messung stattgefunden hat, ist ein wichtiges Beschreibungsmerkmal. Häufige Ursache für die Messung bei höherer Temperatur ist die nicht befriedigende Löslichkeit der Probe in dem Lösungsmittel bei niedrigerer Temperatur.
Neben der Verbesserung der Löslichkeit hat die Temperaturerhöhung in einigen Fällen eine Änderung in der Anzahl der Peaks zur Folge: Durch die Temperaturerhöhung wird die Aktivierungsenergie für Konformationsumwandlungen aufgebracht, das Molekül kann sich in ein anderes Isomer umwandeln. Um diesen Vorgang genauer untersuchen zu können, benötigt man eine ganze Reihe von Messungen, um den Koaleszenzpunkt der sich ineinander umwandelnden Isomeren bestimmen zu können. Die Dokumentation diese temperaturabhängigen NMR-Spektren ist in CHEMIS nicht vorgesehen, da man zu ihrer vollständigen Auswertung ein Strukturformeldarstellungsmodul brauchte.

Die Messungen beim NMR finden in Lösung statt. Also benötigt man Lösungsmittel, die keine Eigensignale liefern. Es kommen Stoffe in Frage, die entweder

- keine Protonen enthalten, wie die perhalogenierten Kohlenwasserstoffe oder
- deren Protonen durch Deuterium ausgetauscht wurden.

Da die perhalogenierten Kohlenwasserstoffe aber oft ein nicht ausreichendes Lösungsvermögen haben, ist man gezwungen, die teureren deuterierten Lösungsmittel zu verwenden. Ähnlich wie bei den Einbettungsmitteln bei der IR-spektroskopie taucht hier die Frage auf, ob man nicht für die in der NMR-Spektroskopie gebräuchliche Lösungsmittel eine besondere Datei eröffnen sollte. Deuterierte Verbindungen stellen in der Stoffdatei Inseln dar, da keine Reaktionen zu ihnen hin führen und sie auch nicht als Ausgangsprodukte für weitere Synthesen verwendet werden. Diese Stoffe gehören nicht in das Beziehungsnetz, das durch die Reaktionsdatei geknüpft wird. Dadurch sind sie genügend hervorgehoben, so daß sich eine besondere Liste erübrigt. Möglich erscheint der Aufbau einer Datei all der Stoffen, die mindestens einmal als Lösungsmittel in der NMR-Spektroskopie Verwendung gefunden haben.
Diese Art der Listenbildung hat einige Vorteile:

- Man vermeidet die zu großzügige Verwendung von Spezialwissen, dessen Auswahl nicht immer willkürfrei möglich ist.
- Es werden nur Listen aufgebaut, die notwendige Informationen enthalten. Datenfriedhöfe werden vermieden.
- Durch die unmittelbaren Informationsbedürfnisse, die zum Aufbau der Listen Anlaß gegeben haben, sind die Listen entsprechend den Anforderungen der Praxis benutzernah aufgebaut. Es wird der Aufbau von wenig gefragten Listen vermieden.
- Jede Liste stellt in ihrem gegenwärtigen Aufbau und Umfang ein Spiegelbild des Systementwicklungsprozesses dar, die Veränderungen lassen sich durch Vergleich der

gegenwärtigen mit den früheren Listen aufzeigen. Mehr noch gilt dies bei der Anwendung der heuristischen Regeln.

Zu jedem Peak (maximale Anzahl: 1o) existiert ein Datenerfassungsschema, das den Benutzer im Dialog zu den einzelnen Felder führt und von ihm die einzugebenden Daten abfragt.
Als erster Wert eines Peaks wird die Größe der chemischen Verschiebung eingegeben, die Einheit ist ppm, relativ zu Tetramethylsilan. Es wird die δ - Skala verwendet, weil sie am gebräuchlichsten ist. Die Auflösung ist o,1 ppm. Es können auch Bereiche angegeben werden, so daß breite Multipletts ebenfalls darstellbar sind. Dabei ist zu beachten, daß die Signale in der numerischen Reihenfolge eingespeichert werden, in der sie auch bei der Ausgabe erscheinen sollen. Es erfolgt intern keine Umsortierung.
Für den Aussagewert eines Spektrums ist die Angabe der Spin-Spin-Kopplungskonstante wichtig, man ist in der Lage, genauere Zuordnungen zu treffen und die Stereochemie ist aufklärbar.
Die Aufspaltung wird durch die

- Multiplizität
  - o Singulett
  - o Dublett
  - o Triplett und
  - o Multiplett und durch
- die Kopplungskonstante (in Hz) beschrieben.

Es ist programmiert, daß bei der Angabe "s" für Singulett nicht mehr nach der Größe der Kopplungskonstante abgefragt wird, bei der Eingabe "m" für Multiplett wird noch nach der Kopplungskonstante gefragt, da man im Multiplett noch einzelne Peaks unterscheiden kann und Kopplungskonstanten ablesbar sind.

Die Zuordnung der Peaks erfolgt unter Verwendung der Schlüsselzahlen aus der Liste der Interpretationen für die IR-Spektren. Diese Liste wird entsprechend erweitert.
Nachdem ein NMR-Spektrum vollständig eingespeichert wurde, fragt das System den Benutzer, ob er zu dem gleichen Stoff noch ein weiteres Spektrum einspeichern will. (Austausch H/D oder Temperaturänderung)
Wird dies verneint, so kann das nächste Spektrum abgespeichert oder die Eingabe beendet werden.

Beispiel für ein NMR-Spektrum:

```
nmr,v 2201
 2201:  MORPHOLIN-4-CARBONSAEURE-(4)-ANILID-(2-MORPHOLINO-CHINAZOLYL-(4)
-IMID)
        N M R - S P E K T R U M  (60 MHz):
        STANDARD:      TETRAMETHYLSILAN
        TEMPERATUR:    30 GRAD C
        LOESUNGSM.:    DEUTEROCHLOROFORM
            ppm        H    STRUKTUR    J (Hz)    ZUORDNUNG
           12,0        1    SINGULETT     -       sekundaeres AMIN
            8,2        1    DUBLETT       -       AROMAT
         6,9   7,5     8    MULTIPLETT    -       AROMAT
            3,7        8    SINGULETT     -       HETEROCYCLUS
            3.7        8    MULTIPLETT    -       HETEROCYCLUS
```

Beispiel für ein NMR-Spektrum mit Austausch:

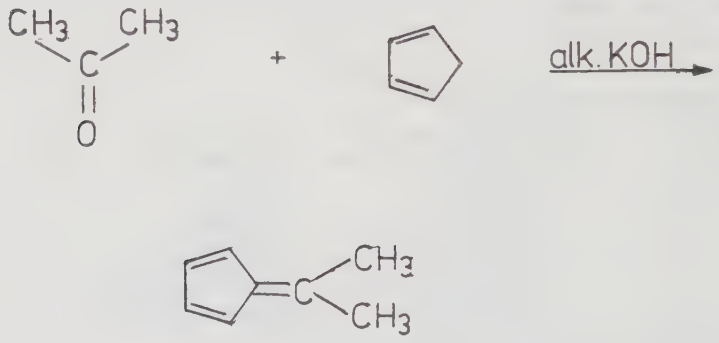
```
nmr,v 2202
 2202:  MORPHOLIN-4-CARBONSAEURE-(4)-(2-METHYLANILID)-(2-MORPHOLINO-8-ME
·HYLCH
        INAZOLYL-(4)-IMID)
        1 .  N M R - S P E K T R U M  (60 MHz):
        STANDARD:      TETRAMETHYLSILAN
        TEMPERATUR:    30 GRAD C
        LOESUNGSM.:    DEUTEROCHLOROFORM
            ppm        H    STRUKTUR    J (Hz)    ZUORDNUNG
           11,2        1    SINGULETT     -       sekundaeres AMIN
            8,1        1    DUBLETT       -       AROMAT
         6,8   7,6     6    MULTIPLETT    -       AROMAT
            3,7        8    SINGULETT     -       HETEROCYCLUS
            3,5        8    MULTIPLETT    -       HETEROCYCLUS
            2,5        3    SINGULETT     -       METHYLGRUPPE
            2,2        3    SINGULETT     -       METHYLGRUPPE
        2 .  N M R - S P E K T R U M  (60 MHz):
        STANDARD:      TETRAMETHYLSILAN
        TEMPERATUR:    30 GRAD C
        LOESUNGSM.:    DEUTEROCHLOROFORM
                       DEUTERIUMOXID
            ppm        H    STRUKTUR    J (Hz)    ZUORDNUNG
            8,1        1    DUBLETT       -       AROMAT
         6,8   7,6     6    MULTIPLETT    -       AROMAT
            3,7        8    SINGULETT     -       HETEROCYCLUS
         3,4   3,8     8    MULTIPLETT    -       HETEROCYCLUS
            2,5        3    SINGULETT     -       METHYLGRUPPE
            2,2        3    SINGULETT     -       METHYLGRUPPE
```

4.4.3. Reaktionsgleichungen

Nachdem in dem vorstehenden Kapitel dargestellt wurde, wie die Daten zu den einzelnen Verbindungen im System behandelt werden, soll nun erläutert werden, wie die Beziehungen zwischen den Stoffen beschrieben, erfaßt und dargestellt werden.
Untersucht man die traditionelle Beschreibung von Reaktionen in der Literatur, so fällt auf, daß neben der Information über das Verhalten der Verbindungen zusätzliche Informationen vermittelt werden:

- strukturelle Informationen
- Daten über die Reaktionsbedingungen
- Ausbeuten
- Literaturzitate etc.

Beispiel :



An dem Beispiel kann man die Bedingungen für die Addition an Ringe studieren und die notwendige Information für die Stereochemie ableiten. CHEMIS verzichtet auf diese strukturelle Information, eine chemische Verbindung wird nicht in funktionelle Gruppe und Gerüst zerlegt, sondern es wird das Verhalten der Verbindung als Ganzes untersucht. Die Beschreibung der Verbindung mit einer Nummer, die keinerlei Auskunft über strukturelle Besonderheiten oder Bildungsweisen vermittelt, macht einen anderen Ansatz zur Beschreibung der chemischen Reaktion notwendig.
Dieses Verfahren macht aber auch einen anderen Ähnlichkeitsbegriff möglich: bei der Formulierung von Ähnlichkeitskriterien ist man nicht mehr auf den Vergleich struktureller Untereinheiten angewiesen, sondern man kann das gesamte Verhalten der chemischen Verbindung betrachten.
Um diese Gegebenheiten nutzen zu können, ist es notwendig, die Beschreibung der Reaktionen darauf abzustimmen.
Im einfachsten Fall besteht die Beschreibung einer Reaktion in einer Vorschrift, die besagt, daß die Codes auf der einen Seite des Trennzeichens die Ausgangsstoffe sein sollen. und daß die Codes auf der anderen Seite des Trennzeichens das oder die Produkte sein sollen. Das Trennzeichen kann man als Operator definieren, der die Edukte in die Produkte überführt. Mit dieser Minimalbeschreibung kann man zwar das Beziehungsnetz zwischen den Stoffen aufbauen, nicht aber jede darüber hinaus gehende Information über die Umwandlung der Stoffe beschreiben. Deshalb werden in der vorliegenden Arbeit Reaktionen als mehrstellige Relationen zwischen Stoffen aufgefaßt.
Daten über die

- Stöchiometrie
- die erzielbaren Ausbeuten
- den Mechanismus
- die Temperatur, bei der die Reaktion abläuft
- den Druck
- das Lösungsmittel und über
- die Katalysatoren

vervollständigen die Beschreibung der chemischen Reaktionen in CHEMIS.
Diese Daten werden aber über ihre deskriptive Funktion hinaus dazu verwendet, um die Reaktionen zu bewerten, sie in Klassen einzuteilen und dadurch dem Chemiker Material für seine Analogieschlüsse an die Hand zu geben.
Jede Reaktion wird mit einer fortlaufenden Nummer gekennzeichnet.
Die Edukte werden mit ihren Verbindungsnummern verschlüsselt, die Anzahl der reagierenden Moleküle wird, durch Kommata getrennt, dahinter eingegeben. Bis zu 10 Datenpaare (Edukt, Anzahl) können eingespeichert werden.
Bei den Produkten erfolgt die Eingabe nach dem gleichen Muster, zusätzlich wird noch nach der Ausbeute gefragt, die, wenn bekannt, ebenfalls abgespeichert werden kann. Ebenfalls 10 Eduktmoleküle sind möglich. (Diese Anzahl reicht in der allermeisten Zahl der Reaktionen aus. Sollte sie wider Erwarten einmal nicht ausreichen, so kann man sich behelfen, indem man sich Supermoleküle definiert, die aus mehreren kleinen Molekülen bestehen.)
Beispiel: Bei der Zersetzung von Glycerintrinitrat in Gegenwart von molekularem Sauerstoff entstehen mehr als 10 Produktmoleküle, man faßt die Stickstoffatome zu einem "Supermolekül" zusammen, ebenso die Moleküle Kohlendioxid, die bei der Reaktion entstehen.

Die Reaktionsart dient dazu, dem Anfrager einen Hinweis auf den Reaktionsmechanismus zu liefern.

Vorschlag für ein Erfassungsschema für Reaktionsdaten:

exec chrein				
Reaktionsnummer:				
Input-Partner / Anzahl der Moleküle:				
Output-Partner / Anzahl der Moleküle / Ausbeute:				
Reaktionsart:				
Temperatur:				
Druck:		atm		Torr
Lösungsmittel:				
Katalysator:				
				ie

Die Liste der Reaktionsarten und ihre Verschlüsselung wird hier maschinenschriftlich wiedergegeben, da diese Liste noch nicht in ihrer Gesamtheit abgefragt werden kann.

Liste der Reaktionsarten

00 = Nicht-Reaktion
01 = radikalische Substitution
02 = nucleophile Substitution am gesättigten C-Atom
03 = Hydrierung, Reduktion
04 = Dehydrierung, Oxidation
05 = Addition
06 = Disproportionierung
07 = Eliminierung
08 = Addition-Eliminierung
09 = Kondensation
10 = Umlagerung
11 = Isomerisierung
12 = Substitution
13 = Addition, Substitution
14 = elektrophile aromatische Substitution
15 = nucleophile Substitution am ungesättigten C-Atom
16 = Substitution mit anschließender Cyclisierung
.
.
.
98 = Reaktionsart wird später festgelegt
99 = Reaktion mit unbestimmten Ergebnis

Diese Liste erfüllt nicht die Anforderungen eines Klassifikationssystems für Reaktionsmechanismen, sie dient dazu, die vorhandenen Reaktionen nach pragmatischen Gesichtspunkten zu ordnen.

Die Temperatur oder der Temperaturbereich, bei dem eine Reaktion stattfindet, wird in ^{o}C eingegeben.
Analog wie bei Druck bei dem Siedepunkt von reinen Stoffen sind bei der Angabe der Reaktionsdrucks Werteangaben sowohl in Atmosphären als auch in Torr erlaubt, die Einheiten werden vom System intern umgerechnet.
Das bei der Reaktion verwendete Lösungsmittel wird mit seiner Verbindungsnummer verschlüsselt, die Angabe von mehreren Stoffen, die ein Lösungsmittelgemisch bilden, ist ebenfalls erlaubt. Im Feld für Lösungsmittel kann auch der Begriff "Schmelze" stehen, wenn die Reaktion ohne die Verwendung von Lösungsmitteln durchgeführt wurde.

Die Katalysatoren werden einer besonderen Liste entnommen. Diese Liste wurde aufgebaut, um auch Gemische und Präparationsformen handhaben zu können, die nicht stöchiometrisch zusammengesetzt sind.

Liste der Katalysatoren:

kat	
1	ZINKCHLORID
2	ZINKOXID/CHROM-(III)-OXID
3	ALUMINIUMPHOSPHAT
4	LUFT/SILBER
5	NICKEL
6	SAEURE
7	KONZ. SCHWEFELSAEURE
8	PLATIN
9	LAUGE
10	ZINNPHOSPHAT
11	LICHT
12	MOLYBDAEN/AKTIVKOHLE
13	AKTIVKOHLE
14	AMMONIAK
15	ALUMINIUMOXID
16	CHROM-(III)-OXID/ALUMINIUMOXID
17	NATRIUMMETHYLAT
18	MOLYBDAEN-(VI)-OXID/ALUMINIUMOXID
19	PALLADIUM/SILICAGEL
20	SILBER
21	QUECKSILBER-(II)-CHLORID
22	VERD. SALZSAEURE
23	PALLADIUM-(II)-CHLORID/KUPFER-(II)-CHLORID
24	KOBALT
25	NICKELTETRACARBONYL
26	KALIUMHYDROXID
27	KUPFER-(I)-CHLORID/AMMONIUMCHLORID/SALZSAEURE
28	QUECKSILBER-(II)-SULFAT/SCHWEFELSAEURE
29	KUPFERCARBID
30	PHOSPHORSAEURE
31	ZINKACETAT/AKTIVKOHLE
32	MANGAN-(II)-ACETAT
33	ALUMINIUMAETHYLAT
34	PALLADIUM/BARIUMSULFAT
35	ALUMINIUMISOPROPYLAT
36	KIESELGEL (ALKALISCH)
37	TANTAL-(V)-OXID/SILICIUMDIOXID
38	VERD. SCHWEFELSAEURE
39	ACETOBACTER
40	NATRIUMHYDROXID
41	EISEN-(II)-SULFAT
42	OSMIUM-(VIII)-OXID
43	KUPFER
44	NATRIUMAMALGAM
45	CHROM/NICKEL
46	NATRIUMALKOHOLAT
47	BORFLUORID
48	KUPFER-/KOBALTACETAT
49	TRIAETHYLPHOSPHAT
50	PYRIDIN
51	KUPFER-(II)-OXID
52	KALIUMHYDROGENSULFAT

Liste der Katalysatoren: (Fortsetzung)

53 WOLFRAM-(VI)-OXID
54 NATRIUMPOLYPHOSPHAT
55 ESSIGSAEURE
56 PIPERIDINACETAT
57 CALCIUM-OXID
58 ZINK/SALZSAEURE
59 NATRIUMHYDROXID/CALCIUMOXID
60 ENZYM
61 VANADINPENTOXID
62 EISEN
63 QUECKSILBER
64 BORFLUORID/WASSER
65 NATRIUMACETAT
66 NATRIUM
67 ZINKOXID
68 WASSERSTOFFPEROXID
69 ROTER PHOSPHOR
70 PALLADIUM
71 PHOSPHORPENTOXID
72 FL. AMMONIAK
73 NICKEL-(II)-SULFAT
74 ACETANHYDRID
75 NATRIUMAMALGAM/ESSIGSAEURE
76 THORIUMDIOXID
77 BARIUMOXID/CALCIUMOXID
78 RESORCIN
79 ASCARIDOL
80 ALUMINIUMCHLORID
81 KUPFER-(II)-OXID/SALZSAEURE
82 KUPFER/CHROM
83 CHROM
84 CALCIUMCHLORID/SALZSAEURE
85 QUECKSILBER-(II)-OXID/BORFLUORIDAETHERAT
86 FLUOBORSAEURE
87 SULFURYLCHLORID
89 BAKTERIEN
90 ANTIMONPENTACHLORID
91 NICKEL/KIESELGUR
92 KUPFER/CHROM-(III)-OXID
93 VANADINPENTOXID/ALUMINIUMOXID
94 KUPFER/KONZ. SCHWEFELSAEURE
95 ALUMINIUMCHLORID/KIESELGEL
96 KONZ. SALZSAEURE
97 FLUORWASSERSTOFF
98 EISEN-(III)-CHLORID
99 EISEN-(III)-BROMID
100 RANEY-NICKEL
101 PLATIN-(II)-CHLORID
102 KOBALT/THORIUM/KUPFER
103 PYRIDIN/PIPERIDIN
104 TRIETHYLAMIN/PYRIDIN

Beispiele für die Darstellung der Reaktionsgleichungen:

```
g,r 1677
 1677:  1114(1) + 2376(1)  =  2286(1)
      1114:  CYANAMID
      2376:  TRICHLORACETONITRIL
      2286:  2.2.2-TRICHLOR-N-CYANACETAMIDIN
     AUSB.:   2286  =  78,0 %
     TEMP.:   0 - 20 GRAD C
     DRUCK:   1 ATM
     LOES.:   AETHANOL
      KAT.:   NATRIUMMETHYLAT
     R-ART:   ADDITION


g,r 1681
 1681:  1607(1) + 1774(2) + 2286(2)  =  1608(1) + 2289(1) + 2298(1)
      1607:  TRIAETHYLAMIN
      1774:  N-p-CHLOR-PHENYL-MORPHOLIN-4-CARBIMIDOYLCHLORID
      2286:  2.2.2-TRICHLOR-N-CYANACETAMIDIN
      1608:  TRIAETHYLAMINHYDROCHLORID
      2289:  2.2.2-TRICHLOR-N-(6-CHLOR-2-MORPHOLINO-4-CHINAZOLINYL)ACE
MIDIN-
             HYDROCHLORID
      2298:  1-(4-CHLORPHENYL)-2-IMINO-6-MORPHOLINO-4-(TRICHLORMETHYL)
.2-DIH
             YDRO-1.3.5-TRIAZIN
     AUSB.:   2289  =  19,0 %
              2298  =   4,0 %
     TEMP.:   20 GRAD C
     DRUCK:   1 ATM
     LOES.:   DIMETHOXYETHAN
     R-ART:   98


g,r 1626
 1626:  32(2) + 2179(1)  =  3(4) + 2182(1)
       32:  THIONYLCHLORID
     2179:  PHENYLEN-1.4-DIACRYLSAEURE
        3:  WASSER
     2182:  3.7-DICHLORBENZO(1.2-b:4.5-b')DITHIOPHEN-2.6-DICARBONYLCH
RID
     AUSB.:   2182  =  41,0 %
     TEMP.:   140 GRAD C
     DRUCK:   1 ATM
      KAT.:   TRIETHYLAMIN/PYRIDIN
     R-ART:   98
      LIT.:   NOTIZ UEBER DIE SYNTHESE VON BEZODITHIOPHENEN
              W. RIED, R. GURYN, G. OREMEK
              LIEBIGS ANN. CHEM. 1980, 1172-1175
```

"Just as manipulation of number in arithmetic is the fundamental mode of operation in contemporary computers, manipulation of symbols in formal systems is likely to be the fundamental operating mode of more sophisticated problemsolving computers of the future."

H. Gelernter, 1959

5. Die Heuristiken

Unter heuristischen Problemlösungsverfahren werden Wege verstanden, die dazu dienen, ein Problem zu vereinfachen, es zu strukturieren, in Untereinheiten zu zerlegen oder es mittels Lösungsansätzen, die mehr allgemeinen Charakter haben, näherungsweise zu lösen.(122)
Man geht dabei von der Annahme aus, daß sich das zu lösende Problem in Einzelprobleme auflösen lasse, die - zumindest in 1.Näherung - voneinander unabhängige Lösungen zulassen. Weiter nimmt man an, daß sich die Lösungen der Einzelprobleme zu einer Gesamtlösung des Ursprungsproblems zusammenführen lassen. Für Probleme aus den Naturwissenschaften, deren Umgebung sich durch Parameter genau beschreiben läßt, ist dieses Verfahren erfolgversprechend. Es wird ein Parameter - unter Konstanthaltung aller anderen - variiert und die Auswirkungen auf das Gesamtproblem studiert.

Beispiel: Eine Reaktion gelingt wider Erwarten nicht. (Man hat, da kein Anlaß bei der Planung zu erkennen war, die gleichen Reaktionsbedingungen wie bei der analogen Reaktion vorher gewählt.) Als erstes wird man die Reaktionsdauer verlängern, insbesondere dann, wenn man dünnschichtchromatographisch einen kleinen Umsatz der Edukte feststellen kann. Hat auch eine verlängerte Reaktionszeit keinen Erfolg, so wird man versuchen, die Reaktionstemperatur zu erhöhen. Man geht dabei schrittweise vor: zuerst wird um ca. 2o - 3o ^{o}C erwärmt und als zweiter Schritt wird bis zum Siedepunkt des Lösungsmittels erhitzt. Bleibt auch dieser Schritt erfolglos, so geht man zu einem ähnlichen, aber höhersiedenden Lösungsmittel über (Diethylether wird durch Dioxan ausgetauscht).

Als weitere Parameter stehen zur Verfügung:

- Verwendung von Katalysatoren
- Aktivierung von Ausgangsstoffen
- Druckerhöhung

und schließlich

- Umplanung der Reaktion oder der gesamten Synthese:
  - andere Ausgangsstoffe
  - Umwege durch Übersynthesen
  - photochemische Verfahren und
  - neue Synthesekonzepte.

Dieses stufenweise Vorgehen beim Lösen der Probleme ermöglicht die Simulation des Problemlösungsverhaltens mittels des Computers.
Ein Teil des Problemlösungsverfahrens besteht in der Beschaffung der Informationen und ihre Auswertung im Rahmen der Fragestellung.
Das Suchen von Informationen, das bei der Konzipierung von Alternativen notwendig ist, kann eine Maschine schneller und zielgerichteter als ein Mensch durchführen, allerdings ist sie nicht in der Lage, andere als die gerade abgefragten Daten zu finden und nachzuweisen. Der menschliche Rechercheur ist fähig, die Fragestellung abzuwandeln, nach Ähnlichkeiten zu suchen oder neue, sich während der Suche ergebende Gesichtspunkte in die Suche mit einzubeziehen.
Der Maschine fehlt diese Flexibilität, sie ist dem Menschen allerdings dann überlegen, wenn es gilt, Routineoperationen mit großer Zuverlässigkeit durchzuführen.
Bei der Planung eines Informationssystems für Chemiker muß man die unterschiedlichen Leistungen des menschlichen Problemlösers und der maschinellen Informationsverarbeitung so miteinander kombinieren, daß sich beide Vorgehensweisen ergänzen. (123)

Die Methoden der heuristischen Programmierung haben zum Ziel, die Denkweisen des menschlichen Problemlösers mittels eines Computers zu simulieren. Die dabei verwendeten Methoden gehören in das Arbeitsgebiet der künstlichen Intelligenz.
Unabhängig von dem, was immer man unter Intelligenz im einzelnen verstehen kann, haben umfangreiche Bemühungen stattgefunden, Leistungen, zu deren Ausführung unstreitig intelligentes Verhalten erforderlich ist, mittels Computern erfolgreich zu simulieren.(124)
Dazu zählen

- Systeme zur Verarbeitung natürlicher Sprache
- mathematische Beweiser
- Deduktionssysteme
- Verfahren zur Mustererkennung und zur Bildverarbeitung und
- Arbeiten auf dem Gebiet der Robotertechnologie.

Im engen Zusammenhang stehen die Arbeiten auf dem Gebiet der Expertensysteme, die versuchen, die Arbeitsweise von z.B. Psychotherapeuten unter engen Voraussetzungen zu simulieren.
Versucht man, CHEMIS in eine dieser Kategorien einzuord- (126)
nen, so stellt CHEMIS ein Expertensystem dar, da es Aufgaben erfüllen soll, die bisher menschlichen Spezialisten vorbehalten waren, wie z.B.:

- die Zuordnung von Spektraldaten zu Verbindungen, die Interpretation von Spektren,(125)
- die Generierung von Ähnlichkeitslisten,
- automatische Klassenbildung aufgrund von vorgegebenen Kriterien,
- Erkennen von Eigenschaften,

und dadurch

- Unterstützung der Problemlösungsarbeit des Chemikers.

(127)

Dieses Verfahren erlaubt die vergleichende Beurteilung alternativer Vorschläge, die vom System geliefert werden. Dabei spielt die Anzahl der Stufen, über die die Synthese läuft, keine Rolle. Unerheblich ist weiter, ob es eine aufbauende oder abbauende Synthese ist.
Sind Reaktionen in der Kette betroffen, für die sich keine Ausbeuteangaben im System finden, so werden sie bei der Berechnung der Gesamtausbeute nicht berücksichtigt. Der Anfrager wird so nicht in die Irre geführt, da ihm die Einzelausbeuten der Schritte nachgewiesen werden, so daß er den Gang der Ausbeuteuerte im Lauf der Synthese nachvollziehen kann.

- Weiter kann der Benutzer nach Kombinationen von Merkmalen fragen: er kann sich einen Weg nachweisen lassen, der
  eine möglichst hohe Ausbeute hat,
  möglichst wenig Stufen erfordert und
  eine vorgegebene Temperatur nicht übersteigt.

Er kann fordern, daß die Reaktion unter milden Bedingungen abläuft. (Temperaturintervall: - 78 °C - + 150 °C; Athmosphärendruck)
Er kann fordern, daß die Synthese über ganz bestimmte Stoffe gehen soll oder daß sie festgelegte Stoffe nicht berührt.
Mit diesen Kombinationsmöglichkeiten bei der Abfrage hat man, ausgehend von einem Merkmal, ausreichende Beschreibungsmerkmale für die Beurteilung einer vom System vorgeschlagenen Reaktionsfolge. Die dabei verwendeten Informationsverarbeitungsprozesse und Bewertungen sind heuristischer Natur, die Ergebnisse dienen dazu, die Entscheidung des Chemikers zu unterstützen. Die Informationen werden zusammengezogen und verdichtet (Berechnung der Gesamtausbeute) oder mittels vorgegebener Kriterien (Vorliegen eines bestimmten Temperaturintervalls) ausgewählt oder verworfen.

Die angestrebte Unterstützungsfunktion des Systems wird durch Heuristiken geleistet. Die Heuristiken greifen auf das gespeicherte Wissen zu und verarbeiten es.
Dabei gilt: Je spezifischer und genauer eine Regel formuliert ist, desto enger wird ihr Gültigkeitsbereich, desto sicherer werden aber die durch die Regel gemachten Aussagen. (128)

Schema:

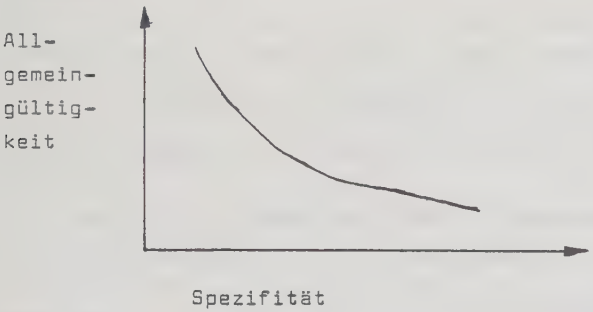


Abb. 22

Das Bemühen bei der Formulierung der Heuristiken geht dahin, möglichst spezifische Heuristiken zu finden, damit die Regeln die im System gespeicherten Daten sinnvoll nutzen können. Die Spezifität und die Sicherheit der Aussagen werden mit einem engen Anwendungsbereich erkauft. Dieser Weg ist angebracht, da man einen möglichst geringen Unterschied zwischen der Verfahrensweise der Maschine und den Heuristiken anstrebt.
Aussagen wie " Bei der Planung von Synthesen sind Reaktionen mit möglichst hohen Ausbeuten anzuwenden. "
sind trivial, sie erhalten einen gewissen Wert erst dann, wenn sie mit weiteren Aussagen verstärkend oder auch abschwächend verbunden werden:

- Die Ausbeuteangaben einer Reaktion werden verwendet, um die Gesamtausbeute einer Reaktionskette zu bestimmen.

Heuristische Regeln können auch Sortiercharakter haben: Wenn bei einer Substanz ein Eigenschaftsspektrum vorliegt, (beschrieben durch physikalische Daten oder durch chemische Reaktionen) so wird sie in eine neu zu bildende Klasse von Stoffen eingeordnet, die eine neue, abstraktere Eigenschaft haben. Diese neue Eigenschaft kann verbal beschrieben werden (hydrolyseempfindliches Chlorformamidin) oder numerisch. (Temperaturintervall - 78 °C - + 150 °C)

Allgemeine Form einer Regel:
Wenn bei einer Substanz S_1 die Eigenschaft E_1 vorliegt, dann ist die Substanz in die Klasse K_1 einzuordnen.

Für die Aussage in negativer Form stehen zwei Alternativen zur Auswahl:
Wenn bei einer Substanz S_2 die Eigenschaft E_1 nicht vorliegt, dann ist die Substanz S_2 nicht in die Klasse K_1 einzuordnen, sondern bleibt unberücksichtigt.

oder:
Wenn bei einer Substanz S_2 die Eigenschaft E_1 nicht vorliegt, dann wird die Substanz in die Klasse K_2 eingeordnet.

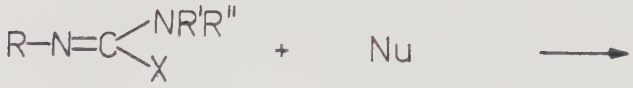
Aus den Einordnungsregeln folgt, daß die Klassen K_1 und K_2 keine gemeinsame Elemente haben. Sie sind einander komplementär. Nun gibt es aber Stoffe, die nicht in eine der beiden gebildeten Klassen einzuordnen sind, sei es, weil über sie noch keine Informationen vorliegen oder weil die vorhandenen Informationen die geforderten Kriterien nicht erfüllen. Diese Stoffe werden in eine besondere Liste eingetragen, so daß man nachsehen kann, welche Wirkungen die Heuristik in ihrer gegenwärtigen Form auf das Sortierergebnis hat, ob das Verfahren überhaupt greift oder ob eventuell nur Scheingruppen gebildet werden.

Die Gesamtheit der gespeicherten Stoffe wird einer ja/nein-Entscheidung unterworfen, die oft in dieser Schärfe nicht zu treffen ist. Auch die Definition weiterer Klassen schafft nur einen bedingten Ausweg, da man Definitionsarbeit leisten muß, um die Einteilung zu verfeinern. Mit jeder neuen Klasse, die definiert wird, bleiben Objekte übrig, die in keine Klasse passen wollen. Die Verfeinerungsarbeit kann ohne Ende weitergeführt werden.
Wir stoßen hier auf den Gegensatz zwischen der strengen algorithmischen Verfahrensweise der Datenverarbeitung und der unscharfen Suche nach Lösungen, die für die heuristische Denkweise charakteristisch ist. Die vom System aufgestellten Listen kann man daher nur als vorläufige Vorschläge betrachten, an deren Verbesserung zwei Mechanismen wirken:

- mit steigender Anzahl der im System eingespeicherten Verbindungen lassen sich bessere Unterscheidungen treffen. Es steht dann mehr Datenmaterial zur Verfügung.
- die Regeln werden laufend verfeinert. Neue Regeln werden aufgestellt und auf ihre Programmierfähigkeit untersucht.

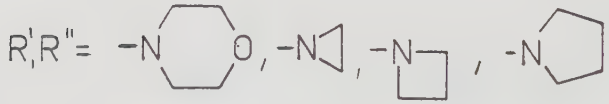
Diese Vorgehensweise wird bei der Untersuchung der Reaktivität der Chlorformamidine gegenüber Nucleophilen untersucht.

5.1. Die Reaktionen der Chlorformamidine

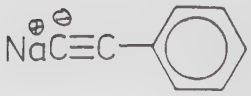


R = Aliphat
Aromat
substituierter Aromat
Heteroaromat

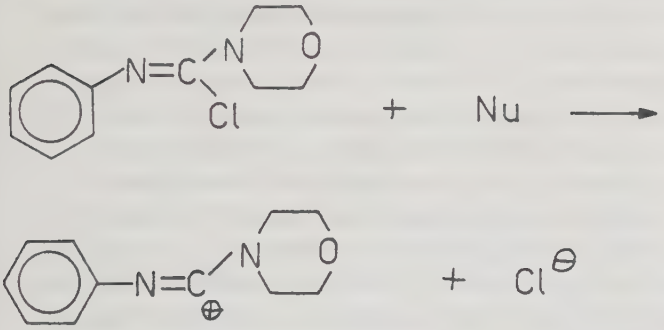
X = F, Cl, Br, I



Nu = H_2O, R_2O, ROH, H_2S, RSH, NH_3, NHR_2,
Br-Mg-Cl, KSCN, AgNCO, KCN, $AgNO_3$



Aus diesen Einzeltatsachen läßt sich dieser Grundreaktionsmechanismus ableiten:



(129, 13o)

Das Primärsubstitutionsprodukt ist nur in einigen Fällen zu fassen, es schließen sich Folgereaktionen an:

- elektrocyclische Reaktionen,
- Dimerisierungen und
- Umlagerungen.

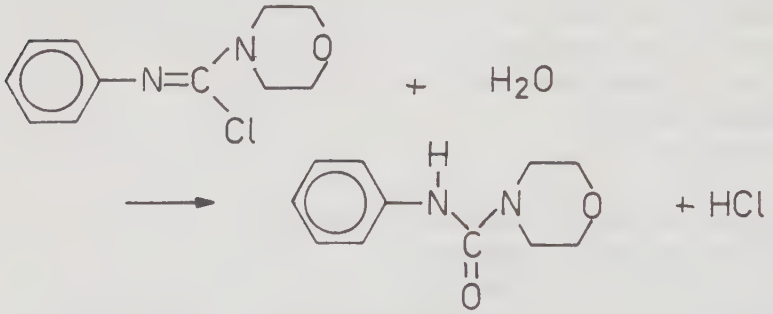
Bei der Betrachtung des Mechanismus treten Fragen auf:

- Kann man das Kation nachweisen ?
- Ist das primäre Substitutionsprodukt überhaupt isolierbar ?
- Welche Weiterreaktionen schließen sich an ?
- Welche Nukleophile reagieren ?
- Welche Nukleophile reagieren nicht ?
- Kann man die Reaktion stufenweise ablaufen lassen, so daß sich alle mechanistisch wichtigen Zwischenstufen als Verbindungen isolieren lassen ?
- Wie sieht die Stereochemie aus ?

Zur Untersuchung dieser Fragen bei möglichst vielen Chlorformamidinen sollte ein Nukleophil herangezogen werden, das mit möglichst vielen Chlorformamidinen reagiert hat oder reagieren kann.

Dieses Nucleophil ist das Wasser, da es mit allen Chlorformamidinen zu den entsprechenden Harnstoffen reagiert. Die Harnstoffe sind gut kristallisierende Produkte, die sich daher leicht charakterisieren lassen, die IR-Spektren sind ebenfalls aussagekräftig.
Darüberhinaus hat die Reaktion mit Wasser noch eine praktische Bedeutung: da Wasser in der Luft als Wasserdampf enthalten ist, liefert eine Untersuchung der Reaktionsfähigkeit der Chlorformamidine gegenüber Wasser Aussagen über die Hydrolyseempfindlichkeit der Stoffe.

Weiterreaktionen, die das Bild komplizieren könnten, treten nicht auf:



Ein HCl-Fänger ist nicht notwendig.
Am Anfang der Untersuchungen wurden nur hydrolyseempfindliche Chlorformamidine betrachtet. Dieser Zugang zu der Verbindungsklasse läßt sich historisch erkären.
Der Begriff der Hydrolyseempfindlichkeit bedarf einer genaueren Aufschlüsselung, da er eine doppelte Bedeutung hat:

- einmal dient er zum Vergleich der Reaktivitäten im Vergleich zu einem bekannten Standard und
- zum zweiten liefert die Betrachtung wichtige Hinweise für die Planung und Durchführung der Synthesen.

Stufen der Hydrolyseempfindlichkeit der Chlorformamidine:

- Substanz muß mit einem Gemisch aus Dimethoxiethan als Lösungsvermittler, Wasser und KOH eine halbe Stunde unter Rückfluß gekocht werden.
- Es genügt, die Stoffe in Wasser aufzuschlämmen und einige Zeit bei Raumtemperatur zu rühren.
- Beim Stehenlassen an der Luft beginnt die Hydrolyse der Substanz.
- Beim Filtrieren der Reaktionslösung zersetzt sich die Substanz.
- Um die Substanz unzersetzt handhaben zu können, ist es erforderlich, die Lösungen und die Trockensubstanz unter Schutzgas zu handhaben.

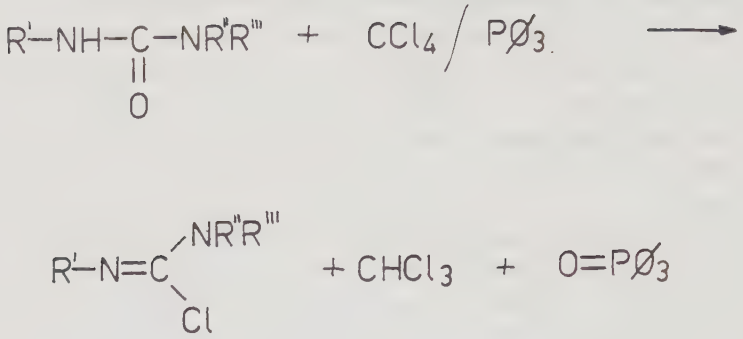
Da die Reaktivität gegenüber dem Modellnucleophil Wasser parallel mit der Reaktivität gegenüber anderen Nucleophilen geht, kann man die mit Wasser gewonnenen Ergebnisse auf die anderen Nucleophile übertragen.
Die erste Einteilung der Chlorformamidine wird in zwei Klassen vorgenommen:

- in die hydrolyseunempfindlichen
  und
- in die hydrolyseempfindlchen Stoffe.

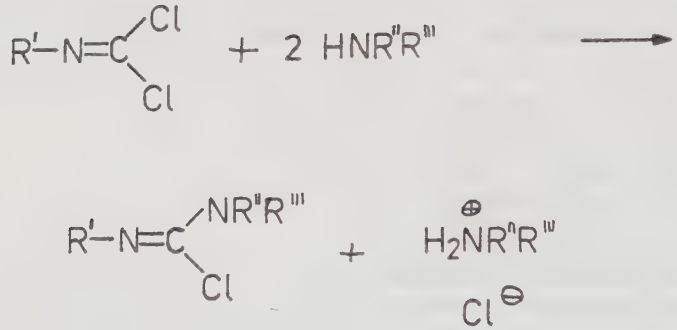
Kriterien für die Klasseneinteilung:
In die Klasse der hydrolyseempfindlichen Chlorformamidine werden alle Stoffe eingeordnet, die

- durch die Appel-Synthese hergestellt werden
  und
- die durch die Umsetzung von Isocyaniddichloriden und sekundären Aminen hergestellt werden.

Reaktion 1:



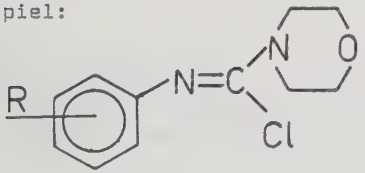
Reaktion 2:



R^1 ist ein aromatischer Rest, die dann synthetisierten Chlorformamidine sind kristallin. Wenn R^1 ein aliphatischer Rest ist, sind die dargestellten Chlorformamidine häufig flüssig, die Öle lassen sich auf Grund der Hydrolyseempfindlichkeit nur schwer reinigen, bei der Destillation sind Verluste nicht zu vermeiden. Man wird daher versuchen, die Synthese so zu planen, daß nach Möglichkeit feste Substanzen entstehen.

Bedingt durch die Größe der Destillationsapparatur und durch die größeren Verluste durch thermische Zersetzung sind bei der Reinigung durch Destillation die Verluste größer als bei der Umkristallisation. (Ausnahme: Kugelrohrdestillation, bei diesem Verfahren ist aber die Trennwirkung nicht sehr groß.) Bei der Umkristallisation erhält man gleichzeitig Informationen über das Löslichkeitsverhalten der Substanz und über das geeignete Laufmittel für die Dünnschichtchromatographie.
Ist es nicht möglich, die Edukte so zu wählen, daß die Produkte fest sind, so ist entweder die Ansatzgröße so zu planen, daß genügend Substanz für die anschließende Destillation zur Verfügung steht oder man kann versuchen, durch Salzbildung (Hydrochloride, Pikrate) ein festes Produkt zu erhalten, das man anschließend leicht reinigen kann. Dieser Weg erfordert aber zwei zusätzliche Reaktionen.
Ein andererer Weg zur Darstellung der festen Produkte besteht darin, möglichst substituierte Produkte anzustreben.

Beispiel:



R	H	4-OEt	2.6-Cl	2.4.6-CH_3	
Fp.	60	88-89	84-86	75	°C

Variiert man die Reste am sekundären Amin, so besagt die Erfahrung, daß die Chlorformamidine besonders dann gut kristallisieren, wenn Morpholin verwendet wird.

Damit kann man folgende Regel formulieren:

Heuristik 1
Wenn keine anderen Gründe dagegen sprechen, wird man versuchen, die Ausgangsstoffe so zu wählen, daß feste Produkte bei der Synthese entstehen.

Es fällt auf, daß Morpholin außerordentlich häufig als Aminrest verwendet wird. Folgende Gründe sprechen dafür:

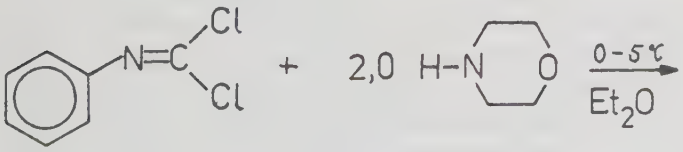
- Morpholin ist eine starke Base. (pK_B = 8,33) Es ist also sowohl für die Appel-Synthese, für die Synthese aus Isocyaniddichloriden als auch für die Synthese nach Bredereck brauchbar.
- Morpholin ist flüssig. Es ist daher einfach zu dosieren und die Reinigung durch Destillation bei Atmosphärendruck stellt kein Problem dar.
- Morpholin ist stabil, eine einmal hergestellte und gereinigte Menge kann portionsweise eingesetzt werden. Die umständliche Destillation unmittelbar vor der Reaktion kann entfallen.
- Morpholin ist preiswert.
- Morpholin ist im IR-Spektrum einer Substanz einfach zu erkennen: bei ca. 1100 cm^{-1} tritt eine scharfe und starke Absorptionsbande auf, die auf die C-O-Valenzschwingung des cyclischen Ethers zurückzuführen ist. Diese Bande ist ein sicherer Hinweis auf das Vorliegen von Morpholin in der Substanz.
- Im NMR-Spektrum kann man die 8 Wasserstoffatomkerne des Morpholinrests als Eichgruppe verwenden, um zu entscheiden, ob eine Reaktion unter Dimerisierung der Ausgangsstoffe stattgefunden hat oder nicht.

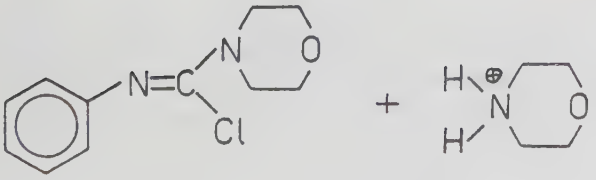
Heuristik 2
Wenn immer möglich, sollte als sekundärer Aminrest Morpholin verwendet werden.

Bei der Synthese von Chlorformamidinen aus den Isocyaniddichloriden und sekundären Aminen ist darauf zu achten, daß das Amin genau doppelt molar zugesetzt wird, um die unerwünschte Nebenreaktion zu den Guanidinen zu vermeiden. Ein temporärer Überschuß ist ebenfalls zu vermeiden. Die Temperatur der Reaktionslösung sollte immer zwischen 0 °C und maximal + 5 °C liegen, steigt die Temperatur über diesen Wert hinaus, so gewinnt die unerwünschte Nebenreaktion zu den Guanidinen die Oberhand.

Heuristik 3
Bei der Synthese der Chlorformamidine aus sekundären Aminen und Isocyaniddichloriden gilt, daß
- ein Überschuß an Amin zu vermeiden ist und daß
- die Temperatur + 5 °C nicht übersteigen darf, um die Bildung von Guanidinen zu verhindern.

Reaktionsschema :





Das sekundäre Amin sollte nukleophil genug sein, um ein Cl-Atom des Isocyaniddichlorids substituieren zu können. Diese Forderung hat zwei Konsequenzen:

- Man benötigt ein Maß für die Nukleophilie der Amine, um die Substanzen miteinander vergleichen zu können.
- Es wird eine Liste gefordert, die die Amine und andere Basen in eine numerische Reihe einordnet.

Nun ist die Nukleophilie einer Substanz nicht direkt meßbar, sondern nur durch mittelbare Überlegungen zugänglich. Es existiert keine Maßeinheit und daher auch keine Maßzahl, die man so ohne weiteres vorhandenen Tabellenwerken entnehmen könnte. Man ist gezwungen, eine gut meßbare physikalische Eigenschaft zu finden, die der Nukleophilie parallel läuft. Bei der Verwendung einer Ersatzgröße für eine gesuchte Eigenschaft geht man das Risiko ein, eine Eigenschaft zu wählen, die eine unbefriedigende Ähnlichkeit mit der unzugänglichen Eigenschaft aufweist. Das ist auch dann der Fall, wenn man eine physikalische Eigenschaft mit einer Beschreibung eines chemischen Verhaltens korrelieren möchte. Chemisches Verhalten kann man nicht isoliert betrachten, es hängt von vielen Einflüssen ab, die ausgewählt werden können.
Die gesuchte Eigenschaft, die als Maß für die Nukleophilie dienen könnte, ist die Basizität einer Verbindung, beschrieben durch die Basenkonstante.

Heuristik 4
Basizität und Nukleophilie von sekundären Aminen laufen gleichsinnig. Starke Basen sind auch gute Nukleophile, schwache Basen sind schlechtere Nukleophile.

Die Heuristik 4 greift auf die Liste des Basenkonstanten zu:

pkb

pKb	V-NR.	NAME
12,60	96:	CALCIUMHYDROXID
11,57	96:	CALCIUMHYDROXID
11,29	2251:	AZETIDIN
11,27	1832:	PYRROLIDIN
11,12	1994:	PIPERIDIN
11,01	1607:	TRIAETHYLAMIN
10,96	1878:	DIISOPROPYLAMIN
10,73	1368:	DIMETHYLAMIN
10,71	1425:	n-PROPYLAMIN
10,66	1123:	METHYLAMIN
10,49	1848:	DIETHYLAMIN
10,04	74:	SILBERHYDROXID
9,81	1369:	TRIMETHYLAMIN
9,33	1796:	BENZYLAMIN
9,25	15:	AMMONIAK
9,18	1736:	PHENYLALANIN
8,33	1768:	MORPHOLIN
8,26	2193:	STRYCHNIN
8,23	17:	HYDRAZIN
8,21	2171:	MORPHIN
8,01	2250:	AZIRIDIN
6,03	39:	HYDROXYLAMIN
5,34	2253:	4-METHOXYANILIN
5,25	1702:	PYRIDIN
4,85	1880:	N-METHYLANILIN
4,63	1972:	ANILIN
4,52	2252:	2-METHOXYANILIN
4,23	2058:	3-METHOXYANILIN
3,15	1583:	ALANYLGLYCIN
2,34	1193:	alpha-ALANIN
0,79	1883:	DIPHENYLAMIN
0,63	1238:	ACETAMID
0,10	1112:	HARNSTOFF

Diese Arbeitshypothese dient als Voraussetzung zur Schaffung von entsprechend strukturierten Listen, die die Informationen in der Form enthalten, in der sie benötigt werden. Wenn man sich vom System eine nach steigender Basizität geordnete Liste ausgeben läßt, kann man annehmen, daß ein Stoff, der oben in der Liste steht, weniger nukleophil ist als der Stoff, der in der Liste weiter unten steht.
Die angenommene Parallelität zwischen Nukleophilie und Basizität kann als grobe Annäherung an die Wirklichkeit aufgefaßt werden, da die Basizität einer Verbindung durch die Lage des Protonierungs/Deprotonierungsgleichgewichts beschrieben wird und die Nukleophilie ein Begriff aus der Reaktionsmechanistik ist, der die Geschwindigkeit einer Reaktion beschreibt. Man sollte sich darüber im Klaren sein, daß eine thermodynamische Größe als Maß für eine kinetische Größe nur in erster Näherung taugt. Man sollte sich bei der Verwendung dieser Skala immer bewußt sein, daß sie nur einen Fingerzeig liefert. Die Gefahr der Überinterpretationen einer numerischen Liste ist immer gegeben.
Sterische Argumente spielen bei der Aufstellung der Liste keine Rolle, man findet auch Amine wie Brucin oder Strychnin in der Liste, Amine, die zur Racemattrennung benutzt werden.
Die Stellung in der Liste besagt weiter nichts über die Leichtigkeit, mit der die Reaktion mit ausgewählten Partnern vor sich geht, sie gibt vielmehr Hinweise auf die grundsätzliche Möglichkeit, die Reaktion durchzuführen.

Bei der Reaktion nach Appel braucht man nicht so kritisch die Mengenvehältnisse der Substanzen untereinander zu beachten, die Gefahr einer dadurch ausgelösten Nebenreaktion besteht nicht. Dafür sind bei der Appel-Reaktion die Ansatzgrößen von Einfluß auf die Ausbeute. Es hat sich herausgestellt, daß die optimale Ansatzgröße bei ca. 0,1 Mol liegt.

Heuristik 5
Bei der Appel-Synthese sollte die Ansatzgröße bei 0,1 Mol liegen, um optimale Ausbeuten sicher zu stellen.

Bei der Beschreibung der Reaktionsfähigkeit der Chlorformamidine fällt die Hydrolyseempfindlichkeit der Stoffklasse auf. Die Skala der Hydrolyseempfindlichkeit ist weit. Die zu treffenden Vorsichtsmaßnahmen sind umfangreich:

- die Apparatur ist vorzutrocknen, um das dem Glas anhaftende Wasser zu entfernen,
- es sind absolute Lösungsmittel zu verwenden und
- man kann unter Schutzgas arbeiten.

Die letzte Möglichkeit erfordert den größten Aufwand, hat aber den Vorteil, sicher zu sein. Verwendet man Argon als Schutzgas, so kann man die Gefäße kurzzeitig offen stehen lassen, da das spezifisch schwerere Argon die Luft verdrängt hat und eine Durchmischung verhindert. Preislich bestehen keine ausschlaggebenden Unterschiede zwischen Reinststickstoff und Argon.

Heuristik 6
Argon ist das optimale Schutzgas.

Gestartet wird die Appel-Reaktion durch Vermischung der Komponenten. Das Ende der Reaktion ist nicht so leicht zu erkennen wie bei der Isocyaniddichlorid-Methode, wo die Reaktion mit den Zutropfen der Amin-Lösung beendet ist. Dünnschichtchromatographisch ist die Reaktion nicht zu verfolgen, da die gesuchten Reaktionsprodukte sich zersetzen. Man muß also ein anderes Kriterium für das Ende der Appel-Reaktion verwenden.

Man findet, daß die Reaktionslösung gegen Ende der Reaktion gelb wird. Bei den bifunktionellen Chlorformamidinen wird die Reaktionslösung klar, vorher war die Reaktionslösung trüb, da die bifunktionellen Harnstoffe schwer löslich sind. Damit hat man ein empirisches Zeitmaß für die Dauer der Reaktion, dieses Maß ist besonders einfach. Die Reaktionsdauer ist ein unkritischer Parameter.

Heuristik 7
Die Reaktionszeit ist ein unkritischer Parameter und leicht zu messen.

Die Reaktionsbedingungen werden im System CHEMIS durch
- Temperatur
- Druck

beschrieben.
Es lassen sich zwei Bereiche unterscheiden:
Bedingungen, deren Realisation und Regelung im Labor keine Probleme schafft und Bedingungen, deren Einhaltung im Labor schwierig zu realisieren ist.
Je nach notwendigem Aufwand wird man versuchen, ihn zu umgehen oder ihn zu treiben.
Für jedes untersuchte Labor wird die Grenze, ab der Bedingungen als zu schwierig zu realisieren angesehen werden, bei anderen Werten liegen. Bei der Einteilung der Daten, in solche, die für milde Reaktionsbedingungen stehen und in die, die für drastische Reaktionsbedingungen stehen, wurden die Maßstäbe angelegt, die für die präparative Arbeit im Arbeitskreis von Prof. Dr. W. Ried an der Universität gelten.
Gemäß der Heuristik 7 bleibt die Reaktionsdauer außerhalb der Überlegungen. Zu diskutieren ist der Einfluß von Druck und Temperatur.

Geprüft werden folgende Fragen:

- Wie groß ist der apparative Aufwand ?
- Braucht man zur Vorbereitung und Realisierung verhältnismäßig viel Zeit ? (teilweise ein Widerspruch zur Heuristik 7)
- Reagiert die Verbindung auf eine Temperaturerhöhung mit Zersetzung ?
- Muß man ein anderes Lösungsmittel verwenden ?

Um diese Fragen beantworten zu können, werden Bereiche definiert, die für milde und für drastische Reaktionsbedingungen stehen.

Heuristik 8
Milde Reaktionsbedingungen liegen vor, wenn die Temperatur sich zwischen - 78 °C und + 150 °C bewegt und bei Atmosphärendruck gearbeitet wird.

Drastische Reaktionsbedingungen herrschen bei Temperaturen, die höher als 150 °C sind und dann, wenn der Druck den Atmosphärendruck übersteigt.

Die Obergrenze von 150 °C für die Reaktionstemperatur ergibt sich aus praktischen Gründen:

- gängige Lösungsmittel sieden unterhalb dieser Temperatur,
- höhersiedende Lösungsmittel werden weniger gern gewählt, da man nach der Reaktion Schwierigkeiten beim Abdestillieren im Rotationsverdampfer hat und das Lösungsmittel sich nur schwer vollständig entfernen läßt,
- die allgemein üblichen Wärmeübertragungsmedien setzen ebenfalls diese Temperaturobergrenze, (zumindest für den Dauerbetrieb) höhere Temperaturen lassen sich durch Sand- oder Metallbäder erreichen.

Die Untergrenze der Reaktionstemperatur ergibt sich aus den im Labor üblichen Kältemischungen und Kühlmedien:

- im Eisbad erreicht man Temperaturen um ca. 0 °C - - 5 °C.
- Eis/Kochsalz-Kältemischungen lassen -33 °C erreichen.
- Eine Trockeneis/Aceton-Mischung kühlt bis -78 °C.
- Mit flüssigem Stickstoff erreicht man -19o °C.

Die letzte Temperatur hat nur eine Randbedeutung, sie wird nur dann benutzt, wenn Kühlfallen in der Vakuumtechnik gekühlt werden müssen. Bei der Temperatur des flüssigen Stickstoffs verlaufen die meisten chemischen Reaktionen für praktische Zwecke zu langsam.
Als sinnvolle Untergrenze der Temperatur liegt damit - 78 °C fest.
Alle Reaktionen, die mit einer Druckerhöhung verbunden sind, sind für einen nicht dafür speziell ausgebildeten und ausgerüsteten Chemiker im Labor schwierig durchzuführen.

- Für Versuche im Autoklaven werden besondere Räume benötigt.
- Ansätze im Bombenrohr sind nicht zu kontrollieren,sie neigen zu spontanen Zersetzungen. Benötigt man den Stoff als Ausgangsprodukt für weitere Synthesen, so ist die diskontinuierliche Reaktionsführung hinderlich.

Wenn es irgendwie geht, wird man also bestrebt sein, bei der Planung von Reaktionen ohne Druckerhöhung auszukommen (bei Hydrierungen durch die Verwendung von Katalysatoren).

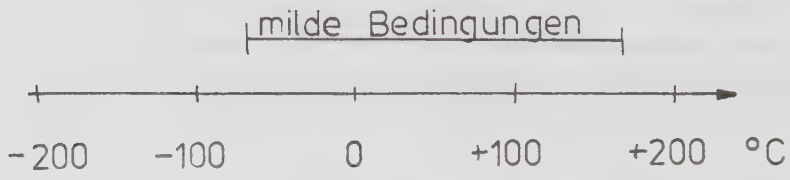


Abb. 23

Wenn man einen Überblick über die im System gespeicherten Reaktionen erhalten möchte, läßt man sich zuerst die Reaktion nachweisen, die unter den oben definierten milden Bedingungen ablaufen, da man bestrebt ist, unnötigen Aufwand zu vermeiden. Die Unterscheidung nach milden und drastischen Reaktionsbedingungen liefert erste Hinweise auf die Eignung einer Reaktion.
Weitere Informationen stecken in den Ausbeutangaben.

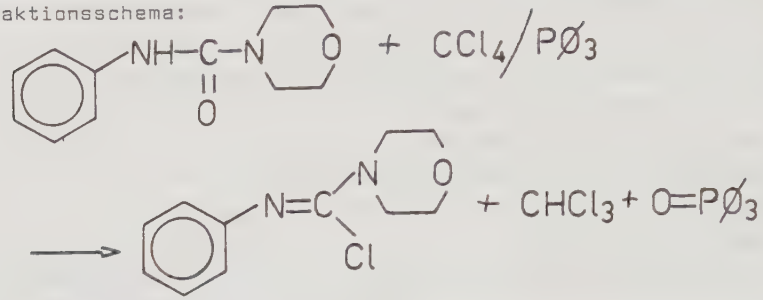
Heuristik 9
Je größer die Ausbeute einer Reaktion, desto günstiger.

Das Kriterium der Ausbeute spielt im Verlauf des Synthesewegs eine zunehmend wichtige Rolle.
Wenn man in den letzten Schritten einer Synthese noch eine relativ geringe Ausbeute in Kauf nehmen kann, so ist es sehr ungünstig, am Anfang einer Reaktionssequenz mit geringen Ausbeuten arbeiten zu müssen. Die Materialverschwendung ist zu hoch.
Wenn man nach Reaktionen fragt, die

- bei milden Bedingungnen ablaufen und
- die eine hohe Ausbeute haben, so kann man davon ausgehen, daß man brauchbare Reaktionen nachgewiesen bekommt.

Heuristiken in der Appel-Synthese:

Reaktionsschema:



Heuristik 10

Die Appel-Synthese wird bei Raumtemperatur durchgeführt.

Als Ausgangsstoffe können substituierte Harnstoffe und Thioharnstoffe dienen. Wenn es möglich ist, setzt man Harnstoffe ein, da das Triphenylphosphinsulfid nur schwer aus dem Reaktionsgemisch zu entfernen ist.

Heuristik 11

Das bei der Appel-Synthese entstehende Triphenylphosphinsulfid (Ausgangsstoff: Thioharnstoffe) ist nur schwer aus dem Reaktionsgemisch abzutrennen. Leichter abtrennbar ist das Triphenylphosphanoxid, daher setzt man bevorzugt Harnstoffe als Ausgangsstoffe ein.

Sowohl Triphenylphosphanoxid als auch Triphenylphosphansulfid sind schwer aus der Reaktionslösung abzutrennen, da sich die beiden Substanzen nur wenig in der Löslichkeit von den Chlorformamidinen unterscheiden. Beim Triphenylphosphansulfid kommt erschwerend die geringe Beständigkeit hinzu, so daß man bei den fälligen Trennoperationen nicht nur die Hydrolyseempfindlichkeit der Chlorformamidine beachten muß, sondern auch die Zersetzlichkeit der Nebenprodukte.

Wegen der Hydrolyseempfindlichkeit der Chlorformamidine scheiden alle protischen Lösungsmittel als Reinigungshilfsmittel aus. Man erreicht eine ausreichende Reinheit der Chlorformamidine durch mehrmaliges Umkristallisieren aus abs. Hexan, die dabei eintretende Ausbeuteverringerung muß man in Kauf nehmen.
Das Triphenylphosphinoxid (TPPO) ist unter den Bedingungen der nukleophilen Substitution stabil. Es besteht also die Möglichkeit, auf die Abtrennung nach der Darstellung der Chlorformamidine zu verzichten und in einer Eintopfreaktion die Substitution der Chlorformamidine vorzunehmen. Die Ausbeuten verbessern sich.

Heuristik 12
TPPO ist stabil und stört bei den weiteren Reaktionen der Chlorformamidine nicht. Man kann also auf eine Abtrennung verzichten.

Das Triphenylphosphinsulfid (TPPS) zersetzt sich aber unter den Reaktionsbedingungen der Substitution.

Heuristik 13
TPPS kann man nicht in der Reaktionslösung belassen, man muß es vorher abtrennen und die Ausbeuteverluste in Kauf nehmen.

Da die Appel-Synthese glatt und mit befriedigenden Ausbeuten verläuft, ist sie der häufigste Zugang zu den hydrolyseempfindlichen Chlorformamidinen.
Es bleibt zu klären, wie es mit den Zugänglichkeiten der Edukte bestellt ist.
Wie gut sind die Harnstoffe im Vergleich zu den Isocyanidchloriden zugänglich?

Tab. 4

Harnstoffe	Isocyaniddichloride
stabil	hydrolyseempfindlich
vielfältig substituiert	Auswahl an Susbtituenten ist etwas beschränkt
problemlos zu handhaben	Stoffe sind reizend

Auch die Eigenschaften der Ausgangsprodukte erklären mit die universelle Verwendbarkeit der Appel-Synthese.

5.2. Diskussion der Hydrolyseempfindlichkeit der Chlorformamidine

Die Heuristik, die die Hydrolyseempfindlichkeit der Chlorformamdine konstatiert, wird im folgenden ausgebaut zu einer Sortierheuristik.
Die unterschiedlichen Grade der Hydrolyseempfindlichkeit wurden schon oben angesprochen. Es bleibt die Aufgabe, Regeln zu finden, die es erlauben, die Chlorformamidine in Klassen einzuteilen, die die Eigenschaften der Verbindungen repräsentieren.
Es ist eine Vorschrift zu formulieren, die die Chlorformamidine in Gruppen sortiert und die die im System gespeicherten Daten verarbeitet. Die Sortierung sollte nach nachvollziehbaren Kriterien erfolgen. Daher wird zum Zwecke der Dokumentation des Sortierprozesses jede erhaltene Stoffliste abgespeichert. Sollte der anfragende Benutzer an der Historie des Entwicklungsprozesses nicht interessiert sein, so kann er die Wiedergabe der Listen unterdrücken.

Die überholten Listen dienen als Arbeitsmittel für den Systementwicklungsprozeß.

1. Versuch

Alle Chlorformamidine, die aus den Harnstoffen und Thioharnstoffen nach der Appel-Synthese dargestellt werden können, sind hydrolyseempfindlich.
Entsprechend sind alle anderen Chlorformamidine - unabhängig vom Syntheseweg - hydrolyseunempfindlich.

Diskussion der Listen:

An der Liste der hydrolyseempfindlichen Chlorformamidine ist nichts auszusetzen, alle darin aufgeführten Stoffe sind hydrolyseempfindlich.
Die Liste der hydrolyseunempfindlichen Chlorformamidine dagegen ist stark verbesserungsbedürftig. Alle Chlorformamidine, die sich sowohl mittels der Appel-Synthese als auch mit der Synthese über die Isocyaniddichloride darstellen lassen, sind ebenfalls in die Liste der hydrolyseempfindlichen Chlorformamidine einzuordnen.

2. Versuch

Alle Chlorformamidine, die nach Appel und die mittels der Isocyaniddichlorid-Methode dargestellt werden können, sind hydrolyseempfindlich.
Alle anderen sind hydrolyseunempfindlich.

Diskussion der Liste:

Die Liste der hydrolyseempfindlichen Chlorformamidine ist umfangreicher geworden, die Fehlerquote ist unverändert gering.

Im Gegensatz dazu ist die Fehlerrate bei den hydrolyseunempfindlichen Chlorformamidine immer noch unbefriedigend hoch.

Zur Verbesserung der Ergebnisse bieten sich diese Auswege an:

- Man verfeinert laufend die Sortierbedingungen, definiert neue Merkmale und Eigenschaften solange, bis sich in keiner der beiden Listen ein Fehler finden läßt.

- Man kann versuchen, sowohl den Inhalt der einen Liste durch eine Reihe von Merkmalen zu beschreiben als auch den Inhalt der weiteren Liste durch Merkmale zu charakterisieren.

Alle Stoffe, die in eine der beiden Listen aufgenommen werden, müssen einen genau definierten Satz von Eigenschaften erfüllen.

Beide Verfahren sind letztendlich unbefriedigend.

Der erste Vorschlag hat zur Folge, daß man gezwungen ist, für jede auch noch so kleine Gruppe von Chlorformamidinen Merkmale und Eigenschaftten zu definieren, die für diese Gruppe spezifisch sind. Im Extremfall findet man in der Klasse nur einen Stoff mit ungewöhnlichen Eigenschaften. Der dafür zu treibende Aufwand wird durch die ansteigende Treffsicherheit der Zuordnungnen nicht gerechtfertigt. Die Grenzen zur intellektuellen Zuordnung werden verwischt, die Sicherheit der Aussagen steigt, es wird aber kein heuristisches Verfahren praktiziert.

Die ausführliche Definition von Listen hat die gleichen Konsequenzen wie die Erhöhung der Anzahl der Listen: sie führen zu einer Zersplitterung der Gesamtheit, ohne das sich der zu treibende Aufwand rechtfertigte.

Es ist also eine Kompromißlösung anzustreben, die sowohl bewirkt,

- daß eine befriedigende Sortierheuristik gefunden wird

als auch

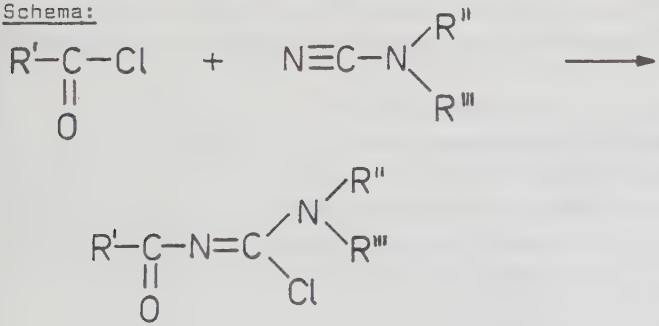
- daß der Arbeitsaufwand nicht prohibitiv hoch wird.

Die Größe des Arbeitsaufwand ist ein wichtiger Gesichtspunkt, da versucht werden soll, die heuristischen Prinzipien auf andere Wissensgebiete zu übertragen, es ist von Nachteil, wenn sich die Regeln in ihrer Struktur als zu kompliziert zum Übertragen erweisen.

3. Versuch

Alle Chlorformamidine, die nach der Appel-Synthese und über die Isocyaniddichloride als Ausgangssubstanzen hergestellt werden können, sind hydrolyseempfindlich.
Alle Chlorformamidine, die nach der Bredereck-Methode darstellbar sind, sind hydrolyseunempfindlich. Ebenso sind die Verbindungen, die aus Sulfensäurechloriden und N-Cyan-Amine hergestellt werden, hydrolyseunempfindlich.

Schema:



Einen unklaren Fall stellen die Hydrochloride der Chlorformamidine dar. Die Hydrochloride selbst werden dargestellt, um die Chlorformamidine zu stabilisieren. Die freien Basen sind wiederum hydrolyseempfindlich.

Die Chlorformamidin-Hydrochloride sind Derivate hydrolyseempfindlicher Stoffe, daher werden sie in der Klasse der hydrolyseempfindlichen Chlorformamidine belassen.
Im System befinden sich noch einige Chlorformamidine, über die keine Daten vorliegen. Eine Zuordnung zu einer der Klassen kann nicht erfolgen, da die Synthesewege, die zu den Stoffen führen, dem System nicht bekannt sind. Die Erkennbarkeit für das System beruht auf der Untersuchung der Reaktionspfade, die zu den Substanzen hin führen. Die verfügbaren physikalischen Daten lassen sich als Ergänzung verwenden.
Die Lage der C=N-Valenzschwingungsbande kann herangezogen werden, um die Grade der Reaktivität einer Klasse genauer zu fassen.
Leider stellt sich heraus, daß das zu Verfügung stehende Datenmaterial eine genaue Einteilung nicht erlaubt, da im Bestand zu große Lücken vorhanden sind. Gerade bei den sehr hydrolyseempfindlichen Chlorformamidinen fehlen naturgemäß die IR-Spektren. Dei vorhandenen Spektren sind oft durch die beginnende Hydrolyse und den Harnstoffanteil nicht eindeutig genug interpretierbar.
Als Merkmale zu Klassifikation werden die Edukte herangezogen.
Die Synthese nach Appel hat diese Edukte:

- Harnstoff oder Thioharnstoff,
- Triphenylphosphan und
- Tetrachlorkohlenstoff.

Als Lösungsmittel wird abs. Dimethoxyethan oder wasserfreies Acetonitril verwendet.
Zur Erkennung der Chlorformamidine wird folgende Abfrage programmiert:

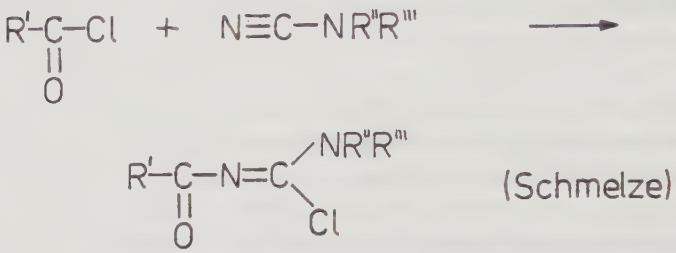
> Immer wenn ein Chlorformamidin (F 61o) aus einem Harnstoff (F6oo) und aus Triphenylphosphan hergestellt wird, dann wird es in die Klasse der hydrolyseempfindlichen Chlorformamidine eingeordnet.

Die Maschine erkennt den Harnstoff an der Familienbezeichnung " F 6oo " und das Triphenylphosphan an der Verbindungsnummer. Chlorformamidine, die aus Isocyaniddichloriden und sekundären Aminen dargestellt werden, kann man durch diese Merkmale erkennen:

- Isocyaniddichloride (F 61o) und
- Amine.

Das Amin kann durch die N-H-Valenzschwingung im Infrarotspektrum erkennen: Interpretationsmerkmal 4o2 sekundäres Amin, Schwingungselement: 12o N-H-Valenzschwingung.
Naheliegender und aussagekräftiger ist es aber, die Basenkonstante der Amine heranzuziehen, da neben der Erkennung der Amine gleichzeitig eine Ordnung der Amine nach steigender Basizität erfolgen kann. Die Dissoziationskonstanten werden einheitlich aus dem "Handbook of Chemistry and Physics" 55^{th} Edition 1974 entnommen.
Es ist immer anzustreben, einen quantitativen Maßstab mit einem qualitativen Merkmal zu verbinden, da die Zusammenführung der Merkmale die Aussagefähigkeit der Skalen steigert.
Als weiteres Merkmal zur Erkennung von hydrolyseempfindlichen Chlorformamidinen kann die Reaktionstemperatur dienen.
Die Appel-Synthese wird bei Raumtemperatur durchgeführt, d.h. es kommt der Temperaturbereich von ca. 15 °C - + 25 °C in Betracht. Bei der Synthese der Chlorformamidine aus Isocyaniddichloriden wird ein Temperaturbereich von 0 °C - +5 °C abgefragt.

Vorschlag für einen Algorithmus zur Erkennung der hydrolyseunempfindlichen Chlorformamidine:
Diese Stoffklasse wurde durch die Synthese nach Bredereck im größeren Maß erschlossen:



Es hat sich herausgestellt, daß die Reaktion nur dann gut durchzuführen ist, wenn man in der Schmelze aus Säurehalogenid und N-Cyan-Amin arbeitet. Lösungen der Stoffe, auf die Reaktionstemperatur gebracht und gerührt, reagieren nicht oder nur sehr langsam. Man erklärt dieses Verhalten durch eine zu stabile Solvatation der Ausgansstoffe durch die Lösungsmittelmoleküle.
Die Mischung der Ausgangsstoffe wird im Trockenschrank auf ca. 80 °C temperiert und reagiert in etwa 8 h vollständig durch. Rühren des Gemisches ist nicht erforderlich.
Die Ausbeuten bei der Synthese sind i.a. gut.
Das eingesetzte N-Cyan-Amin sollte sehr rein sein, ist dies nicht der Fall, so neigt der Reaktionsansatz zu unkontrollierten Zersetzungsreaktionen, der Ansatz geht verloren.
Das N-Cyan-Amin wird durch Umkristallisation gereinigt.
Die Verwendung eines Schutzgases ist nicht erforderlich.
Die Reaktion zeichnet sich durch große Einfachheit in der Durchführung, gute Zugänglichkeit der Edukte und hohe Ausbeuten aus.
Als Carbonylkomponente können alle Carbonsäurehalogenide verwendet werden. Gut gelingt die Reaktion mit ziehend substituierten aromatischen Carbonsäurehalogeniden.

ny,f 610
HYDROLYSE-EMPFINDLICHE CHLORFORMAMIDINE:
(AUSGANGSPRODUKT IST EIN HARNSTOFF ODER EIN ISOCYANIDDICHLORID)
1864: N-(4-CHLORPHENYL)-N'-DIMETHYLCHLORFORMAMIDINHYDROCHLORID
1827: N-4-CHLOR-PHENYL-N'-DIMETHYL-CHLORFORMAMIDN
1774: N-p-CHLOR-PHENYL-MORPHOLIN-4-CARBIMIDOYLCHLORID
1851: (4-CHLORPHENYL)-PYRROLIDIN-1-CARBIMIDOYLCHLORID
1861: CYCLOHEXYL-PYRROLIDIN-1-CARBIMIDOYLCHLORID-HYDROCHLORID
1778: 4-(N-(3.5-DIBROM-4-BIPHENYLYL)-CHLORFORMIMIDOYL)-MORPHOLIN
1818: N.N'-DI- 3-CHLOR-PHENYL)-CHLORFORMAMIDIN-HYDROCHLORID
1877: N-(2.6-DICHLOR-1.3.5-TRIAZIN)-N'-DIISOPROPYLCHLORFORMAMIDIN
1874: N-(2.6-DICHLOR-1.3.5.-TRIAZIN)-N'-DIMETSYL-CHLORFORMAMIDIN
1885: N-(2.6-DICHLOR-1.3.5-TRIAZIN)-N'-DIPHENYLCHLORFORMAMIDIN
1882: N-(2.6-DICHLOR-1.3.5-TRIAZIN)-N'-METHYL-PHENYL-CHLORFORMAMIDIN
1879: N-(2.6-DICHLOR-1.3.5-TRIAZIN)-MORPHOLIN-4-CARBIMIDOYLCHLORID
1788: N.N'-DICYCLOHEXYL-CHLORFORMAMIDIN-HYDROCHLORID
1836: N.N'-DIISOPROPYL-CHLORFORMAMIDIN-HYDROCHLORID
1853: N.N'-DI-(4-METHOXYPHENYL)-CHLORFORMAMIDINHYDROCHLORID
1820: N.N'-DI-(4-METHOXYPHENYL)-CHLORFORMAMIDIN
1815: N.N'-DIMETHYL-CHLORFORMAMIDIN-HYDROCHLORID
1855: N.N'-DI-(beta-NAPHTHYL)-CHLORFORMAMIDINHYDROCHLORID
1839: N.N'-DIOCTADECYL-CHLORFORMAMIDIN-HYDROCHLORID
1845: N.N'-DIPHENYLCHLORFORMAMIDINHYDROCHLORID
1841: N.N'-DIPHENYL-N'-ISOPROPYL-CHLORFORMAMIDIN-HYDROCHLORID
1793: 1.2-DIPHENYL-VINYL-MORPHOLIN-4-CARBIMIDOYLCHLORID
1870: DI-PYRROLIDINOCHLORFORMAMIDINHYDROCHLORID
1806: C-MORPHOLINO-N-(4-ETHOXY-1-NAPHTHYL)-FORMIMIDOYLCHLORID
1804: C-MORPHOLINO-N-(3.4.5-TRIMETHOXYPHENYL)-FORMIMIDOYLCHLORID
1776: N-m-NITRO-PHENYL-MORPHOLIN-4-CARBIMIDOYLCHLORID
1813: N-PHENYL-N'-CYCLOHEXYL-CHLORFORMAMIDIN
1849: N-PHENYL-N'-DIETHYL-CHLORFORMAMIDIN
1862: N-PHENYL-N'-DIMETHYLCHLORFORMAMIDINHYDROCHLORID
1847: N-PHENYL-N'-DIMETHYL-CHLORFORMAMIDIN
1872: N-PHENYL-N-METHYL-N'-DIMETHYLCHLORFORMAMIDINHYDROCHLORID
1769: N-PHENYL-MORPHOLIN-4-CARBIMIDOYLCHLORID
1857: N-PHENYL-PYRROLIDIN-1-CARBIMIDOYLCHLORID-HYDROCHLORID
1868: PYRROLIDIN-1-DIMETHYLCARBIMIDOYLCHLORID-HYDROCHLORID
1784: N-trans-STYRYL-MORPHOLIN-4-CARBIMIDOYLCHLORID
1843: N-p-TOLUOLSULFONYL-N'-n-BUTYLCHLORFORMAMIDIN
1772: N-o-TOLYL-MORPHOLIN-4-CARBIMIDOYLCHLORID

* WASSERFREI ARBEITEN - SCHUTZGAS ARGON *

nhy,f 610

HYDROLYSE-UNEMPFINDLICHE CHLORFORMAMIDINE:
(AUSGANGSPRODUKT IST WEDER EIN HARNSTOFF NOCH EIN ISOCYANIDDICHLORID)

1831: N-BENZYL-CHLORFORMAMIDIN-HYDROCHLORID
1823: BROMFORMAMIDINHYDROBROMID
1829: N-tert.-BUTYL-CHLORFORMAMIDIN-HYDROCHLORID
1916: (3-CHLOR-2-CARBONYLBENZO-(b)-THIOPHENO)-MORPHOLIN-4-CARBIMIDOYLCH
1915: (3-CHLOR-2-CARBONYLBENZO-(b)-THIOPHENO)-PIPERIDIN-1-CARBIMIDOYLCH
1914: (3-CHLOR-2-CARBONYLBENZO-(b)-THIOPHENO)-PYRROLIDIN-1-CARBIMIDOYLC
D
1921: 3-CHLOR-N-(CHLOR(1-PYRROLIDINYL)-METHYLEN)-6-NITRO-BENZO(b)THIOPH
CARBOXAMID
1811: N.N'-DI-n-BUTYLCHLORFORMAMIDINHYDROCHLORID
2062: 3.6-DICHLOR-N-(CHLOR(DIMETHYLAMINO)-METHYLEN)-BENZO(b)THIOPHEN-2-
XAMID
1909: 2.4-DINITROPHENYLSULFENYL-MORPHOLIN-4-CARBIMIDOYLCHLORID
1904: (2.4-DINITRO-PHENYLSULFENYL)-PIPERIDIN-1-CARBIMIDOYLCHLORID
1899: (2.4-DINITRO-PHENYLSULFENYL)-PYRROLIDIN-1-CARBIMIDOYLCHLORID
1809: N.N'-DIPHENYL-CHLORFORMAMIDIN-HYDROCHLORID
1825: N-METHYL-CHLORFORMAMIDIN-HYDROCHLORID
1910: 2-NITRO-4-CHLORPHENYLSULFENYL-MORPHOLIN-4-CARBIMIDOYLCHLORID
1905: (2-NITRO-4-CHLOR-PHENYLSULFENYL)-PIPERIDIN-1-CARBIMIDOYLCHLORID
1900: (2-NITRO-4-CHLOR-PHENYLSULFENYL)-PYRROLIDIN-1-CARBIMIDOYLCHLORID
1911: 2-NITRO-4-METHYLPHENYLSULFENYL-MORPHOLIN-4-CARBIMIDOYLCHLORID
1906: (2-NITRO-4-METHYL-PHENYLSULFENYL)-PIPERIDIN-1-CARBIMIDOYLCHLORID
1901: (2-NITRO-4-METHYL-PHENYLSULFENYL)-PYRROLIDIN-1-CARBIMIDOYLCHLORID
1908: (2-NITRO-PHENYLSULFENYL)-MORPHOLIN-4-CARBIMIDOYLCHLORID
1902: (2-NITRO-PHENYLSULFENYL)-PIPERIDIN-1-CARBIMIDOYLCHLORID
1898: (2-NITRO-PHENYLSULFENYL)-PYRROLIDIN-1-CARBIMIDOYLCHLORID

* KEINE BESONDEREN SCHUTZMASSNAHMEN ERFORDERLICH *

610
FORMAMIDINE OHNE AUSGANGSPRODUKT IM DATENBESTAND:
: AZETIDIN-1-CARBONSAEURE-(3.4-DIOXO-2-CHLOR-CYCLOBUT-1-ENYLIMID)-CHLORID
: AZETIDIN-1-CARBONSAEURE-(3.4-DIOXO-2-PHENYL-CYCLOBUTEN-1-ENYLIMID)-CHLORID
: AZIRIDIN-1-CARBONSAEURE-(3.4-DIOXO-2-PHENYL-CYCLOBUTEN-1-ENYLIMID)-CHLORID
: 4-(N-(4-BIPHENYLYL)-CHLORFORMIMIDOYL)-MORPHOLIN
: N.N'-BIS-(CHLOR-MORPHOLINOMETHYLEN)-p-PHENYLENDIAMIN
: 6-BROM-3-CHLOR-N-(CHLOR(DIMETHYLAMINO)-METHYLEN)-BENZO(b)THIOPHEN-2-CARBOXAMID
: N-(3-CHLOR-2-CARBONYLBENZO-(b)-THIOPHENO)-IMINOKOHLENSAEURE-(2.4-DIMETHYL-PHENYL)-ESTER-CHLORID
: 3-CHLOR-N-(CHLOR(DIMETHYLAMINO)-METHYLEN)-BENZO(b)THIOPHEN-2-CARBOXAMID
: 3-CHLOR-N.N'-bis(CHLOR(DIMETHYLAMINO)METHYLEN)BENZO(b)THIOPHEN-2.6-DICARBOXAMID
: 3-CHLOR-N-(CHLOR(DIMETHYLAMINO)-METHYLEN)-6-CYAN-BENZO(b)THIOPHEN-2-CARBOXAMID
: 3-CHLOR-N-(CHLOR(DIMETHYLAMINO)-METHYLEN)-6-FLUOR-BENZO(b)THIOPHEN-2-CARBOXAMID
3-CHLOR-N-(CHLOR(DIMETHYLAMINO)-METHYLEN)-6-METHYL-BENZO(b)THIOPHEN-2-CARBOXAMID
4-CHLOR-N-(4-CHLOR-2.5-DIMETHYL-PHENYL)-FORMIMIDOYL-MORPHOLIN
3-CHLOR-N-(CHLORMORPHOLINO-METHYLEN)-BENZO(b)THIOPHEN-2-CARBOXAMID
3-CHLOR-N-(CHLOR(1-PYRROLIDINYL)-METHYLEN)-6-METHYL-BENZO(b)THIOPHEN-2-CARBOXAMID
C-CHLOR-N'-(4-ETHOXYPHENYL)-N-METHYL-N-PHENYL-FORMAMIDIN
CHLORFORMAMIDINHYDROCHLORID
C-CHLOR-N'-MESITYL-N-METHYL-N-PHENYL-FORMAMIDIN
C-CHLOR-N'-(2-METHOXYPHENYL)-N-METHYL-N-PHENYL-FORMAMIDIN
C-CHLOR-N'-(3-METHOXYPHENYL)-N-METHYL-N-PHENYL-FORMAMIDIN
C-CHLOR-N'-(4-METHOXYPHENYL)-N-METHYL-N-PHENYL-FORMAMIDIN
1.5-bis-(CHLOR-MORPHOLINO-METHYLENAMINO)-NAPHTHALIN
N.N'-bis-(CHLOR-MORPHOLINO-METHYLEN)-p-PHENYLENDIAMIN
4-(N-(2.6-DICHLOR-PHENYL)-CHLORFORMIMIDOYL)-MORPHOLIN
N-(2.6-DICHLOR-1.3.5-TRIAZIN)-N'-DIETHYLCHLORFORMAMIDIN
N-DIMETHYL-N'-METHYLCHLORFORMAMIDINHYDROCHLORID
4-(N-MESITYL-CHLORFORMIMIDOYL)-MORPHOLIN
N-(6-METHYL-3-CHLOR-2-CARBONYLBENZO-(b)-THIOPHENO)-N'-DIMETHYLCHLORFORMAMIDIN
(6-METHYL-3-CHLOR-2-CARBONYLBENZO-(b)-THIOPHENO)-MORPHOLIN-4-CARBIMIDOYLCHLORID
(6-METHYL-3-CHLOR-2-CARBONYLBENZO-(b)-THIOPHENO)-PYRROLIDIN-1-CARBIMIDOYLCHLORID
N-METHYL-N.N'-DIPHENYL-CHLORFORMAMIDIN
2-METHYL-PHENYL-MORPHOLIN-4-CARBIMIDOYL-CHLORID
MORPHOLIN-4-CARBONSAEURE-(3.4-DIOXO-2-CHLOR-CYCLOBUT-1-ENYLIMID)-CHLORID

...

```
 2110:  MORPHOLIN-4-CARBONSAEURE-(3.4-DIOXO-2-PHENYL-CYCLOBUT-1-ENYLIMID)-
        RID
 1805:  C-MORPHOLINO-N-(4-METHYLMERCAPTOPHENYL)-FORMIMIDOYLCHLORID
 2083:  4-(N-(1-NAPHTHYL)-CHLORFORMIMIDOYL)-MORPHOLIN
 1920:  N-(2-NITROPHENYL)-MORPHOLIN-4-CARBIMIDOYLCHLORID
 2375:  N-(p-PHENYL-PHENYL)-MORPHOLIN-4-CARBIMIDOYLCHLORID
 2122:  PIPERIDIN-1-CARBONSAEURE-(3.4-DIOXO-2-PHENYL-CYCLOBUTEN-1-ENYLIMID
        LORID
 2124:  PYRROLIDIN-1-CARBONSAEURE-(3.4-DIOXO-2-CHLOR-CYCLOBUT-1-ENYLIMID)-
        RID
 2116:  PYRROLIDIN-1-CARBONSAEURE-(3.4-DIOXO-2-PHENYL-CYCLOBUT-1-ENYLIMID)
        MID
 2115:  PYRROLIDIN-1-CARBONSAEURE-(3.4-DIOXO-2-PHENYL-CYCLOBUT-1-ENYLIMID)
        ORID
 2114:  PYRROLIDIN-1-CARBONSAEURE-(3.4-DIOXO-2-PHENYL-CYCLOBUT-1-ENYLIMID)
        ORID
 2117:  PYRROLIDIN-1-CARBONSAEURE-(3.4-DIOXO-2-PHENYL-CYCLOBUT-1-ENYLIMID)
        ID
 2332:  C-PYRROLIDINO-N-(1-CHLOR-3.4-DIOXO-CYCLOBUTENYL)-FORMIMIDOYLCHLORI
 1985:  TETRAETHYLCHLORFORMAMIDINHYDROCHLORID
 1866:  N.N'-TETRAMETHYLCHLORFORMAMIDINHYDROCHLORID
 1787:  3.4.6-TRICHLOR-N-(CHLOR(DIETHYLAMINO)-METHYLEN)-BENZO(b)THIOPHEN-2
        BOXAMID

* DIESE LISTE WIRD SICH DURCH EINGABE VON AUSGANGSSTOFFEN UND *
* REAKTIONSGLEICHUNGEN AENDERN                                *
?
```

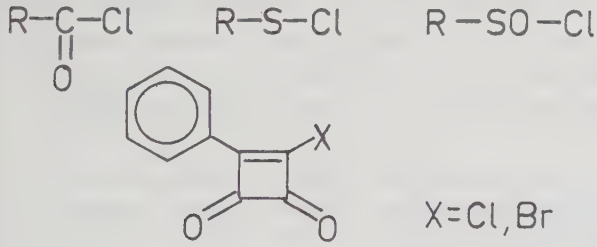
Bei der Bearbeitung der Synthese entstand auch die folgende Musterfrage:

" Ich benötige alle aromatischen Carbosäurehalogenide, die am Kern ziehend substituiert sind und die etwa zwischen 5o °C und 15o °C schmelzen."

CHEMIS kann diese Frage beantworten.

- Die Schmelzbereiche sind abfragbar.
- Ob eine Verbindung aromatische Charakter hat oder nicht, erfährt man aus der Suche nach den Merkmalen im IR-Spektrum.
- Über die ziehende Substitution des aromatischen Kerns gibt das IR-Spektrum ebenfalls Auskunft.

Verbindungstypen, die bei der Bredereck-Synthese eingesetzt werden können:



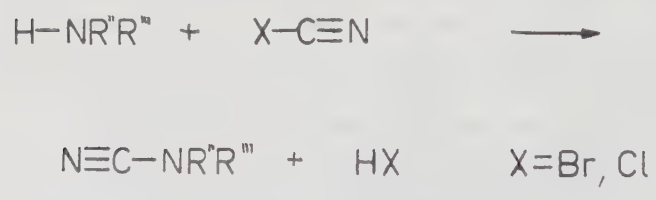
Vorschläge für das Erkennen der Bredereck-Reaktion, die zu hydrolyseunempfindlichen Chlorformamidinen führt:

- über die Acylkomponente
  - o Namen der Stoffe: im chemischen Namen des gesuchten Stoffes muß die Buchstabenfolge "...chlorid..." oder "...bromid..." oder "...chlor..." oder "...brom..." vorkommen. Durch die Maskierungstechnik erreicht man die Wiedergabe der zutreffenden Namen.
  - o IR-Spektrum: die Kohlenstoff-Sauerstoff-Doppelbindung absorbiert im Bereich zwischen 175o und 185o cm^{-1}.

Da aber in diesem Bereich noch viele andere starke Absorptionsbanden liegen, genügt die Suche nach einer Bande allein in diesem Bereich nicht. Die Interpretationsverschlüsselung für die Carbonsäurehalogenidschwingung wird als zusätzliches und eindeutiges Merkmal verwendet. (Schlüsselzahl: 601)
Die Abfrage sähe dann also so aus:

- Zuerst wird die Wiedergabe des IR-Spektrums danach abgesucht, ob sich in der Teilliste der Wellenzahlen ein Eintrag im Bereich zwischen 1750 cm^{-1} und 1850 cm^{-1} vorliegt. Ist eine Bande in diesem Bereich gefunden, so wird gefragt, ob die korrekte Zuordnung vorliegt.
- Die zweite Komponente, die vorliegen muß, ist das N-Cyan-Amin. Erkennbar ist diese Verbindungsklasse an der Buchstaben- und Zeichenfolge "...N-Cyan-Amin..." oder "...N-Cyan-...". Da die Schreibweise der Namen aber nicht einheitlich zu sein braucht, ist dieses Verfahren fehleranfällig, insbesondere dann, wenn man es nicht auf eine Verbindungsklasse mit ausgefeilter Nomenklatur anwendet.

Um diese Schwierigkeiten zu umgehen, wird nach einem eindeutigen Kriterium gesucht, das eine zweifelsfreie Abfrage erlaubt.
Die N-Cyan-Amine werden durch die Umsetzung von sekundären Aminen mit Bromcyan oder Chlorcyan dargestellt:

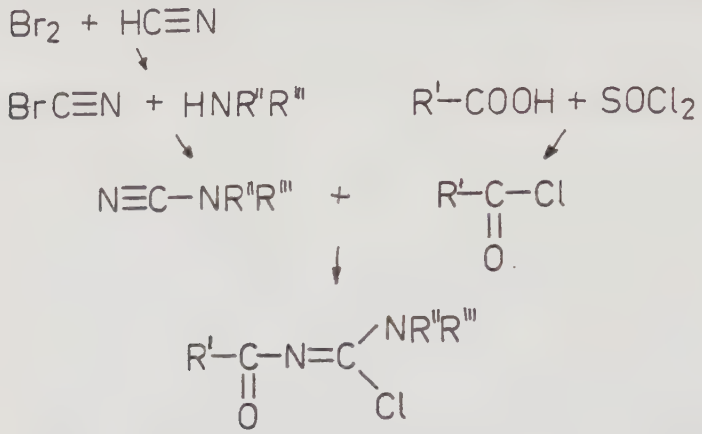


Die Ausbeuten sind bei Chlorcyan etwas besser.
Als Sortierkriterium kann man in diesem Fall den Namen verwenden: "Chlorcyan" und "Bromcyan". Bei den einfachen Verbindungen tauchen keine Probleme auf. Sollten vergleichbare Verbindungen in das System aufgenommen werden, so ist darauf zu achten, daß die Schreibweise der Namen so gewählt wird, daß die Einheitlichkeit gewahrt bleibt.
Von diesen Verbindungen liegen keine IR-Spektren vor, der Weg über die Identifikation der Verbindungen über Absorptionsbanden scheidet aus. Da diese Verbindungen als fertiges Produkt vom Hersteller bezogen werden, stellt das Fehlen der spektralen Information keinen Informationsverlust dar, der schwer wiegt.
Wenn die Frage nach dem Namen des Stoffes kein zufriedenstellendes Ergebnis liefert und auch die Spektren als Informationsquellen nicht zur Verfügung stehen, bleibt noch die Reaktionsbeschreibung des Stoffes für seine Erkennung übrig.

Darstellung von Bromcyan:

(Chlorcyan wird analog dargestellt.)
Um die Hydrolyseunempfindlichkeit eines Chlorformamidins, das nach Bredereck dargestellt wird, eindeutig beschreiben zu können, muß man den Syntheseweg bis zu den einfachen Ausgangsstoffen betrachten.

Stammbaum der hydrolyseunempfindlichen Chlorformamidine:



Als Regel, die dem Algorithmus zu Grunde liegt, kann man formulieren:

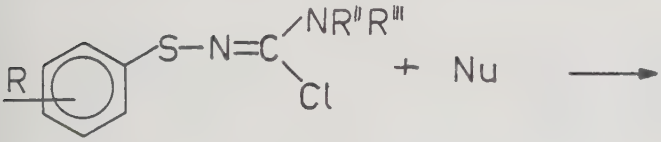
- Chlorformamidine sind dann hydrolyseunempfindlich, wenn sie folgende synthetische Vorstufen haben:
  - o entweder ein Säurehalogenid
    oder
  - o Chlor- oder Bromcyan.

Die Regel gilt auch bei Sulfensäurechloriden und Sulfinsäurechloriden.

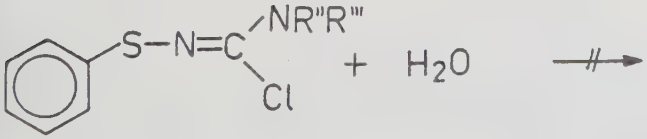
Bei der Aufstellung des Stammbaumes, der zu den hydrolyseunempfindlichen Chlorformamidinen führt, trifft man auf Stoffe, die hydrolyseempfindlich sind: die Säurechloride. Wenn man die Synthese plant, darf man nicht davon ausgehen, daß während der gesamten Synthese Stoffe verwandt werden, die ebensowenig hydrolyseempfindlich sind wie die Zielverbindungen.
Bei den hydrolyseempfindlichen Chlorformamidinen gilt die Verallgemeinerung ebenfalls nicht: sie werden aus hydrolyse-

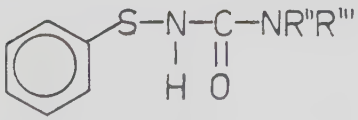
unempfindlichen Harnstoffen synthetisiert.
Aus den Eigenschaften des Produkts kann man nicht umgekehrt auf die Eigenschaften der Edukte schließen.

Es gibt nun Chlorformamidine, die sich nicht in das Schema hydrolyseempfindlich/hydrolyseunempfindlich einordnen lassen. Es sind die Verbindungen, über die zu wenig Daten vorliegen, um eine Einordnung vornehmen zu können.
Es gibt auch Verbindungen, die nicht in der erwarteten Weise unter Substitution des Chloratoms reagieren:

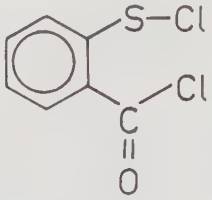


Die Sulfenylharnstoffe sind auf diese Weise nicht darstellbar:

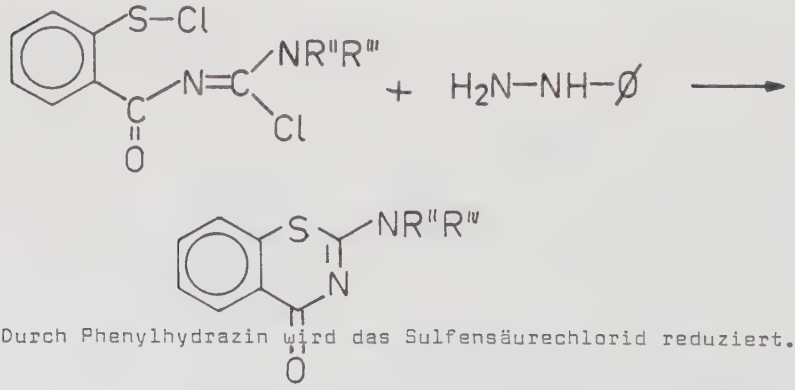




Hauptreaktion ist die Rückbildung der Ausgangsstoffe. Vermutlich liegt es an der geringen Stabilität der Schwefel-Stickstoff-Bindung gegenüber der Schwefel-Schwefel-Bindung.
Interessant wäre es, ein Molekül zu haben, das beide Chlorformamidin-Funktionen in sich vereinigt, da man auf diese Weise die unterschiedlichen Reaktivitäten der beiden Gruppen abschätzen könnte.
2-Chlorcarbonyl-benzolsulfenylchlorid ist eine solche Substanz:

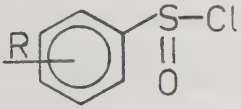


Das Acyl-Chlorformamidin ist nicht nur stabiler, sondern auch reaktiver als das Sulfenyl-Chlorformamidin.
Die geringe Stabilität der S-Cl-Bindung kommt auch in dieser Reaktion zum Ausdruck:



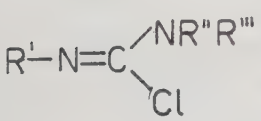
Durch Phenylhydrazin wird das Sulfensäurechlorid reduziert.

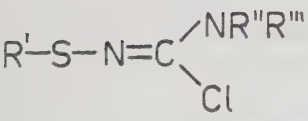
Bei den Sulfinylchlorformamidnen ist die Dimerisierung nicht mehr möglich:

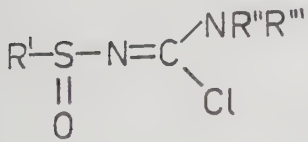


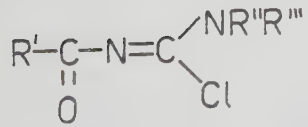
Eine Untersuchung der Eigenschaften steht noch aus.

Die Reaktivität der Chlorformamidne kann in diese Reihe geordnet werden:





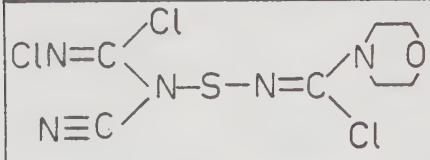
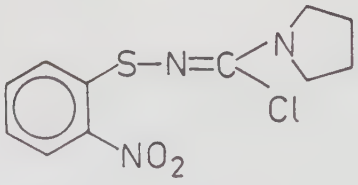




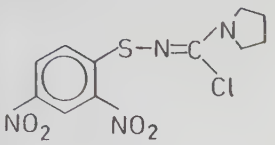
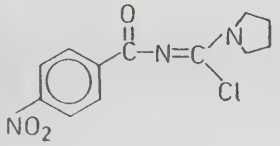
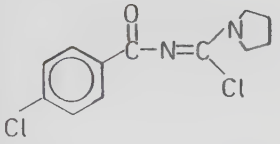
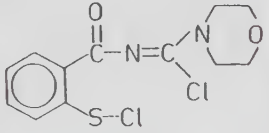
abnehmende Reaktivität gegenüber Nukleophilen

Es erhebt sich nun die Frage nach einer genaueren Zuordnung der Reaktivität der Chlorformamidine als durch das Eigenschaftstripel hydrolyseempfindlich/hydrolyseunempfindlich/nicht zuordenbar. Die unterschiedlichen Grade der Reaktivität müßten sich darstellen und aus Daten ableiten lassen. Da die Reaktivität gegenüber Wasser parallel der Reaktivität gegenüber anderen Nukleophilen verläuft, kann man die Ergebnisse, die man mit H_2O gewonnen hat, versuchsweise auf andere Nukleophile übertragen.
Daher wurde die Hydrolyseempfindlchkeit ausführlich analysiert und maschinell verarbeitbar gemacht.

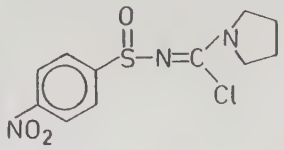
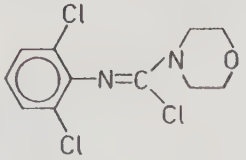
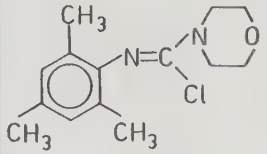
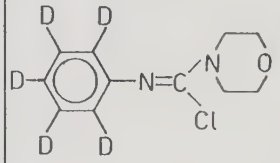
Ein genaueres Merkmal zur Einordnung der allgemeinen Reaktionsfähigkeit vermutet man in der Lage der C-N-Valenzschwingungsbande des Amidin-Systems.
Die Tabelle dient als Beleg für die Hypothese.

Verbindung	Lage der C=N-Schwingung
	1650
	k.A.
	1625
	k.A.
	(cm^{-1})

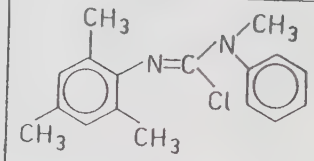
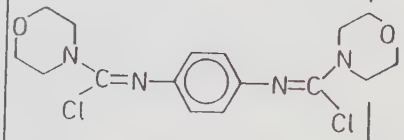
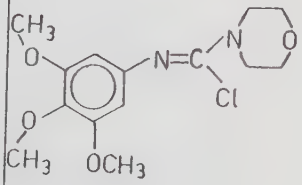
Fortsetzung

Verbindung	Lage der C=N-Schwingung (cm^{-1})	
	1697	k.A.
	1585	—
	ca. 1570	—
	1622	—

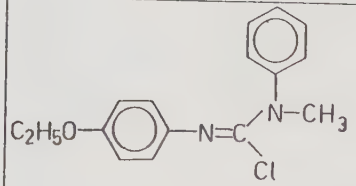
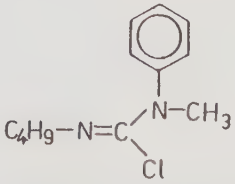
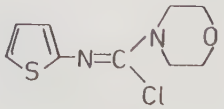
Fortsetzung

Verbindung	Lage der C=N-Schwingung (cm^{-1})	
	1655	k.A.
	1635	+
	1640	+
	1630	+

Fortsetzung

Verbindung	Lage der C=N-Schwingung (cm^{-1})	
	1670	+
	1645	+
	1660	+

Fortsetzung

Verbindung	Lage der C=N-Schwingung (cm^{-1})	
	1655	+
	1640	+
	1645	+

Tab. 5

Erläuterungen zur Tabelle

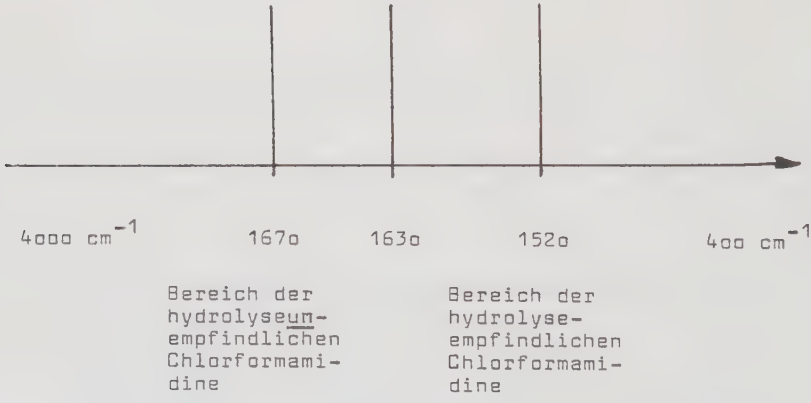
k.A.: Zu diesen Stoffen waren keine weiteren Reaktionsdaten nachgewiesen. Die experimentelle Überprüfung der Reaktionseigenschaften steht noch aus.

\- : Dieses Chlorformamidin ist nicht hydrolyseempfindlich.

\+ : Dieser Stoff ist hydrolyseempfindlich.

Für einige Chlorformamidine liegen keine IR-spektroskopischen Daten vor, die Substanzen sind zu leicht hydrolysierbar. Durch die Hydrolyse der zu messenden Substanz wird der KBr-Preßling trübe und die Transparenz sinkt. Daher liegen zuverlässige Daten in ausreichender Menge nur von den hydrolyseunempfindlichen Chlorformamidinen vor, bei den hydrolyseempfindlichen Chlorformamidinen haben die Messungen mehr stichprobenartigen Charakter.
Die zu formulierende Regel bedarf der weiteren Nachprüfung in Zukunft.
Sie lautet: Chlorformamidine mit einer C=N-Valenzschwingungsbande bei 163o cm^{-1} und größeren Wellenzahlen sind hydrolyseempfindlich.
Chlorformamidine mit einer C=N-Valenzschwingungsbande unterhalb 163o cm^{-1} sind hydrolyseunempfindlich.

Schema:



Die Sulfinyl-Chlorformamidine und die Sulfenyl-Chlorformamidine fallen aus dem Gültigkeitsbereich der Regel heraus, da von ihnen keine Reaktionen mit Wasser, die zu den Harnstoffen führen, bekannt sind.
Der Gedanke an einen Analogieschluß liegt nahe:
Unmittelbare experimentelle Daten liegen nicht vor.
Die Synthese der Sulfenyl-Chlorformamidine erfolgt aus vergleichbaren Komponenten wie die der Acyl-Chlorformamidine.
Die vorhandenen IR-Daten lassen ebenfalls eine Einordnung in die Klasse der hydrolyseunempfindlichen Chlorformamidine gerechtfertigt erscheinen.
Durch Analyse von Verbindungen, deren chemisches Verhalten und deren physikalische Eigenschaften bekannt sind, erhält man Klassifikationsmerkmale, die eine Einordnung von Verbindungen auch dann mit einiger Sicherheit erlauben, wenn deren Datensätze nicht vollständig bekannt sind. Man kann also auch Teileigenschaften vorhersagen.
Dieses Verfahren kann man als einen Versuch ansehen, die Idee des Analogieschlusses so zu erweitern, daß eine entsprechend programmierte Maschine in der Lage ist, ihn nachzuvollziehen. (131)

Bei der Untersuchung der Arbeiten auf dem Chlorformamidingebiet stellte sich heraus, daß sich das chlorierende Reagenz bei der Appel-Synthese in situ auf zweierlei Art und Weise darstellen kann:

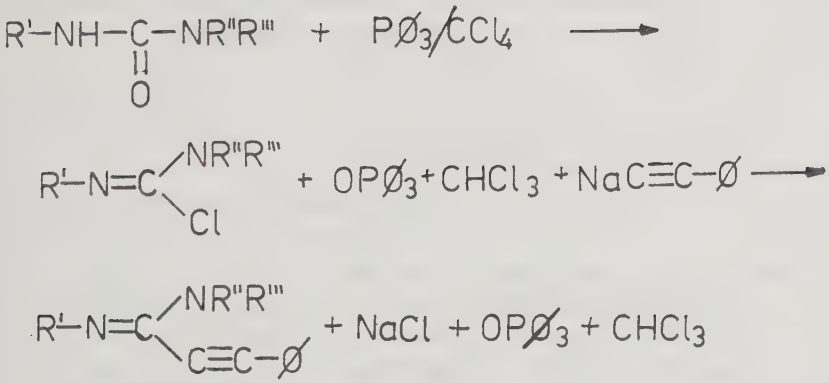
- einmal durch die Reaktion von Triphenylphosphan mit Tetrachlorkohlenstoff
- und zum zweiten durch die Darstellung von Dichlortriphenylphosphoran aus Triphenylphosphan und Phosgen.

Die frisch bereitete Lösung wird zur Synthese eingesetzt.
Die zweite Variante wird aber selten eingesetzt, da der Umgang mit dem giftigen Phosgen nach Möglichkeit vermieden wird. Alle Chlorformamidine, die nach der zweiten Methode

dargestellt wurden, sind auch nach der ersten Variante darstellbar. Daher ist es nicht notwendig, diese Variante der Synthese explizit zu programmieren. Die zweite Variante bietet keine prinzipielle Vorteile.

Nachdem nun die Eigenschaften und die Synthesen der Chlorformamidine behandelt wurden, soll nun auf die Regeln eingegangen werden, die bei der Handhabung zu beachten sind. Der allgemeine Gesichtspunkt der Hydrolyseempfindlichkeit und die daraus sich ergebenden Folgerungen für die Reaktivität gegenüber Nukleophilen wurde schon oben diskutiert. Bei der Reinigung der Chlorformamidine tritt eine Schwierigkeit auf: Das bei der Appel-Synthese entstehende Triphenylphosphanoxid ist mit den bei der Reinigung anwendbaren unpolaren, aprotischen Lösungsmitteln schwer von den Chlorformamidinen zu trennen, da sich die Löslichkeiten stark ähneln. Die säulenchromatographische Auftrennung des Reaktionsansatzes ist nicht möglich, da das notwendige wasserfreie Arbeiten nicht zu realisieren ist. Das Füllmaterial der Säule hat einen Restwassergehalt, beim Einsammeln der Fraktionen kommen diese zwangsläufig mit der Luftfeuchtigkeit in Kontakt.
Man kann eine Reinigung der Produkte durch mehrmaliges Umkristallisieren erreichen, die dabei eintretende Ausbeuteverringerung durch Verschleppung muß man akzeptieren. Diese umständliche Prozedur kann man vermeiden, da das Triphenylphosphanoxid unter den Bedingungen der Substitution des Chloratoms stabil ist. Man nutzt die im Vergleich zu den Chlorformamidinen große Reaktionsträgheit des Triphenylphosphanoxids aus. Das Verfahren wurde bei der Synthese der Propiolamidine ausprobiert. (89)

Reaktionsschema:



Die Abtrennung der stabilen Propiolsäureamidine von den Nebenprodukten bereitet keine Schwierigkeiten. Durch die Einsparung eines Reinigungsschrittes steigt die Ausbeute der Gesamtreaktion an. "Eintopfreaktionen" verfügen immer über diesen Vorteil.

Eine weitere Erhöhung der Löslichkeitsdifferenz zwischen dem Chlorformamidin und dem Triphenylphosphanoxid kann durch die Bildung des Hydrochlorids erreicht werden:

Die Hydrochloride sind im allgemeinen hydrolysebeständige, gut kristallisierende Substanzen.
Auch die Chlorformamidine sind basische Substanzen, die stabile Hydrochloride bilden. Trotz des zusätzlichen Reaktionsschritts geht man diesen Weg, wenn man die Chlorformamidine in eine über längere Zeit stabile Form überführen will.

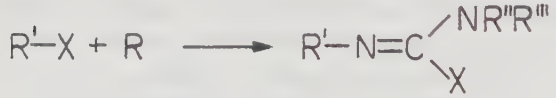
Da die Bildung von Chlorformamidin-Hydrochloriden bei den aliphatischen Chlorformamidinen bevorzugt angewandt wird, wird dieses Verfahren hier nicht weiter diskutiert.

Die homologe Reihe der Chlorformamidine

Bisher wurde untersucht, welche Konsequenzen die Variation der Reste R^1, R^2 und R^3 auf die Eigenschaften der Chlorformamidine hat. Die Untersuchung der Eigenschaften der Fluorformamidine, Bromformamidine und Iodformamidine steht noch aus. Auf Grund der bekannten Reaktivitätsunterschiede der Halogenverbindungen kann man erwarten, daß Brom- und Iodverbindungen reaktiver als die Chlorverbindungen sind. Man sollte daher bestrebt sein, von einer nicht zu reaktiven Klasse der Halogenamidine auszugehen, um nicht zu instabile Verbindungen zu erhalten.

Variante 1

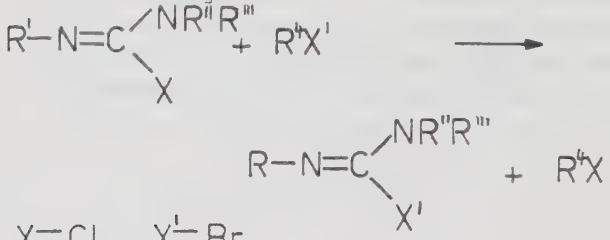
Umsetzung einer halogenierten Verbindung mit einem Partner:



X=F, Cl, Br, I

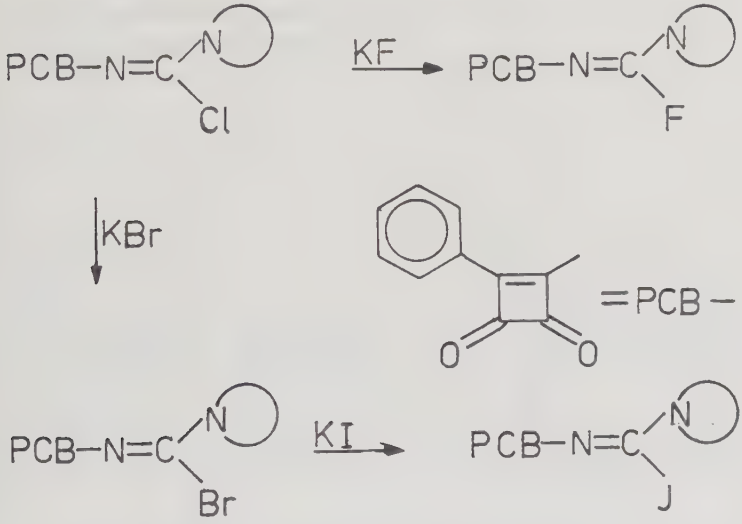
Variante 2

Umsetzung eines Haloformamidins mit einem halogenaustauschendem Reagenz:



X=Cl X'=Br
X=Br X'=J
etc.

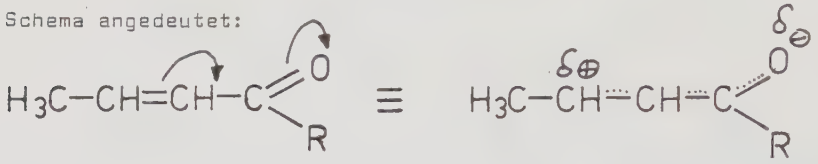
Gewählt wurde die Halogenaustauschmethode, da sich bei diesem Verfahren mehr Variationen anbringen lassen:



Das IR-Spektrum der Halogenamidine zeigt eine starke und breite Bande für die C-N-Valenzschwingung des Amidin-Systems bei 1660 - 1530 cm^{-1}. Der Schwerpunkt der Bande liegt bei ca. 1580 cm^{-1}. Eine Abhängigkeit der Bandenlage vom Halogenatom läßt sich mit den vorliegenden Daten nicht konstatieren. Die Einordnung in die Klasse der hydrolyseunempfindlichen kann über die Lage der C-N-Valenzschwingungsbande erfolgen. Dies ist ein weiteres Beispiel für den oben abgeleiteten Analogieschluß.
Bis auf die Fluorformamidine ist dieser Schluß zutreffend. Es ist eine bekannte Tatsache bei der Aufstellung von homologen Reihen, daß die ersten Glieder (F, B, Be und Li) unregelmäßige Eigenschaften aufweisen. (90)

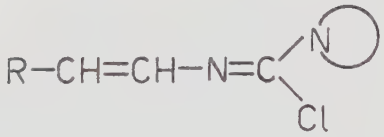
Den Analogiealgorithmus kann man dadurch perfektionieren, indem man eine Abfrage nach den Elementen einbaut, die zu den Unregelmäßigkeiten neigen. Verbindungen, die diese Elemente enthalten, werden bei der Einordnung von Stoffen in Klassen nicht berücksichtigt.

Ein weiteres Prinzip zur Erklärung von Ähnlichkeiten und zur Schaffung von Skalen zur empirischen Beschreibung von Reaktivitäten ist das Viylogieprinzip. (132)
Die Elektronendichte verteilt sich etwa so, wie in dem Schema angedeutet:

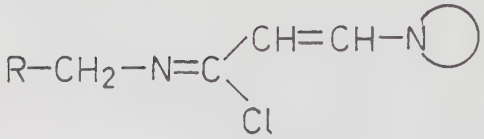


Bei den Chlorformamidinen kann man zwei Vinylogiearten unterscheiden:

Fall 1

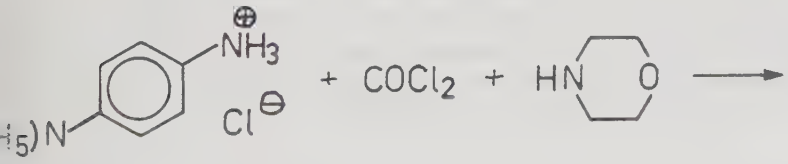


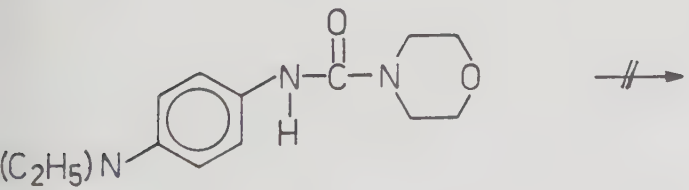
Fall 2

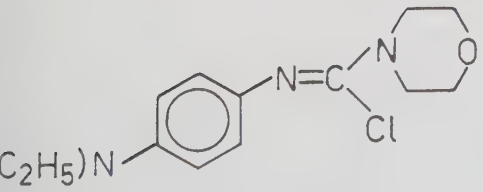


Im Fall 1 besteht die Vinylogie zwischen der funktionellen Gruppe und dem Rest, im allgemeinen dem Aromat. Durch die Elektronenverschiebung werden die Einflüsse der Substituenten weiter geleitet.
Im Fall 2 findet die Verschiebung der Elektronen innerhalb der funktionellen Gruppe statt, die Substituenten haben weniger Einfluß. Im engeren Sinn vinylog ist nur der Fall 2.

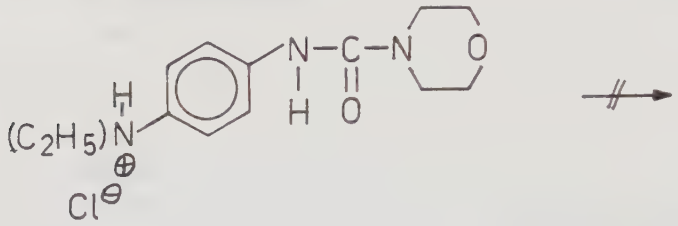
Hypothese:
Chlorformamidine, die am aromatischen Rest R^1 elektronenziehende Substituenten tragen, reagieren langsamer als die Chlorformamidine, die unsubstituiert sind oder die einen elektronenliefernden Rest tragen.
Das kann soweit gehen, daß ein Chlorformamidin, das zweimal schiebend substituiert ist, nicht darstellbar ist:
(89)







Auch die Fixierung des freien Elektronenpaars am Stickstoffatom der Aminogruppe durch Salzbildung führt nicht zum erfolgreichen Syntheseverlauf, die Verminderung des +M-Effekts der Diethylaminogruppe reicht nicht aus,um die Reaktion zu ermöglichen.



Das Chlorformamidin entsteht nicht einmal als instabile, aber abfangbare Zwischenverbindung, die man durch die Weiterreaktion mit einem Nukleophil charakterisieren könnte. (Nukleophil: KSCN oder Lithiumphenylacetylid) Lebensdauerabschätzungen wurden nicht vorgenommen.

Die Erfindung einer Technik, die es ermöglicht, mit der heutigen Anhäufung des Wissens Schritt zu halten, ist zu einem der dringendsten Anliegen der Menschheit geworden.

... Wenn die Wissenschaft Ordnung in das Leben gebracht hat, so muß jetzt Ordnung in die Wissenschaft gebracht werden.

Ortega y Gasset

6. Experimenteller Teil

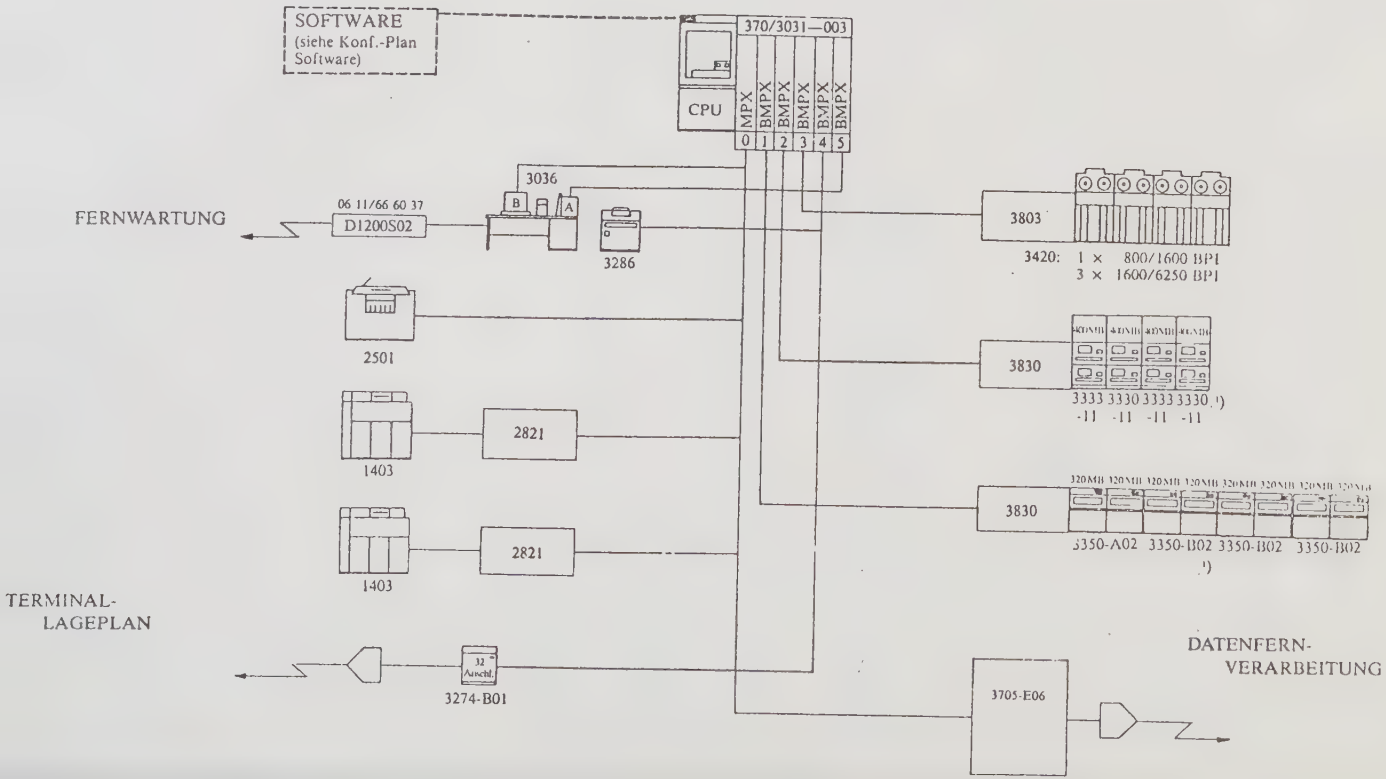
6.1. Der benutzte Computer

Das Informationssystem wurde auf einer IBM-Rechenanlage des Typs /37o-3o31 mit 4MB Hauptspeicherplatz implementiert.
Die Steuerung der Programmabläufe und der Terminals erfolgt unter dem Betriebssystem OS/VS 2 im paging-Modus.
Die Übertragung der Daten erfolgt entweder lokal oder über das DATEX-L Netz der Deutschen Bundespost mit 2.4oo baud nach Heidelberg.
Als Terminals werden die Modelle des Typs IBM 3276 (in Heidelberg) oder IBM 3278 (lokal) verwendet.
Programme und Daten werden auf Magnetplattenspeichern IBM 335o mit einer Kapazität von 32o MB vorgehalten.

Die Programme sind in Assembler geschrieben und haben einen Umfang von ca. 1o.ooo Programmzeilen.
Die Programmierung führte Herr Ketterer vom der GID/SfS in Heidelberg durch.
Der Umfang der Dateien schwankt zwischen 11 Schutzgruppen für Peptide und ca. 65o.ooo Derivat-Sätzen chemischer Verbindungen.

KONFIGURATIONSPLAN — IBM 370/3031

GID (SIT)
Stand: 24.1.81



5.2. Abfrage

In diesem Teil der Arbeit soll die Bedienungsanleitung des Systems CHEMIS ausführlich dargestellt werden, so daß der Benutzer in der Lage ist, den Dialog mit der Maschine zu steuern.
Zur weiteren Vertiefung wird auf das Benutzerhandbuch in der Fassung vom Dezember 1979 verwiesen.

Es werden die Codes für die Abfrage der physikalischen Eigenschaften einzelner Stoffe, für Spektren und Daten zu Reaktionen genau beschrieben und durch Beispiele erläutert.

Die Suchfragen bestehen
- aus dem Code
- aus dem Code und einem odere mehreren Parametern
oder
- aus dem Code, einem oder mehreren Parametern und aus Konstanten.

Diese Parameter beschreiben einen konkreten Meßwert oder einen Meßwertebereich, den der Benutzer sich selost definieren kann.
Die Codes werden unverändert eingegeben.
Die Parameter (dargestellt durch Kleinbuchstaben) werden an die vorgesehenen Stellen eingesetzt.
Die Konstante (dargestellt durch Großbuchstaben) wird unverändert übernommen.
Bei den Abfragen nach numerischen Werten kann man sowohl nach konkreten Einzelwerten als auch nach vorgegebenen Bereichen abfragen.

Die Toleranzen kann man entweder in Form von Ober- und Untergrenzen eines Bereichs eingeben oder in Form einer vorgegebenen Schwankungsbreite um einen konkreten Meßwert. Insbesondere bei der Abfrage nach Spektraldaten erweist sich die Vorgabe einer Toleranz bei Linien oder Peaks als zweckmäßig.

An erster Stelle der Eingabe steht immer das Symbol für die unbekannte Größe, hier sind es alle dem System bekannten Eigenschaften der Verbindung.
An zweiter Stelle steht die bekannte Beschreibung der zu suchenden Größe, hier der Code "V" für die Verbindung, getrennt durch einen Zwischenraum "_" folgt die Verbindungsnummer, der Parameter.
Diese Grundstruktur der Formulierung gilt für alle Abfragen.

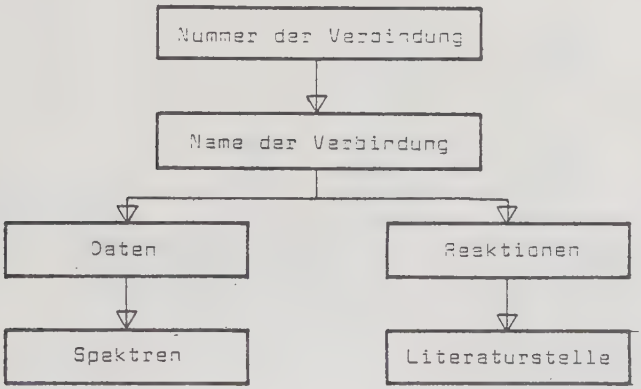
Beispiel:

Ei,V_a

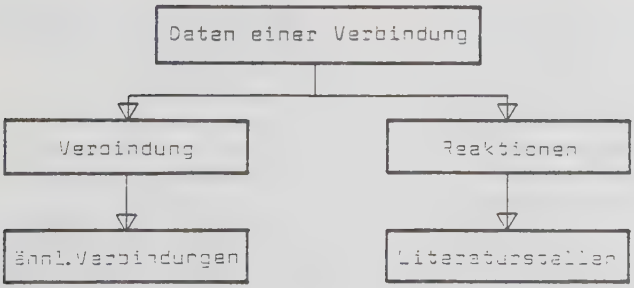
(Abfrage nach allen Eigenschaften einer Verbindung)

Die Flußdiagramme für die Abfragen fassen die Möglichkeiten der Informationswiedergewinnung in CHEMIS zusammen.

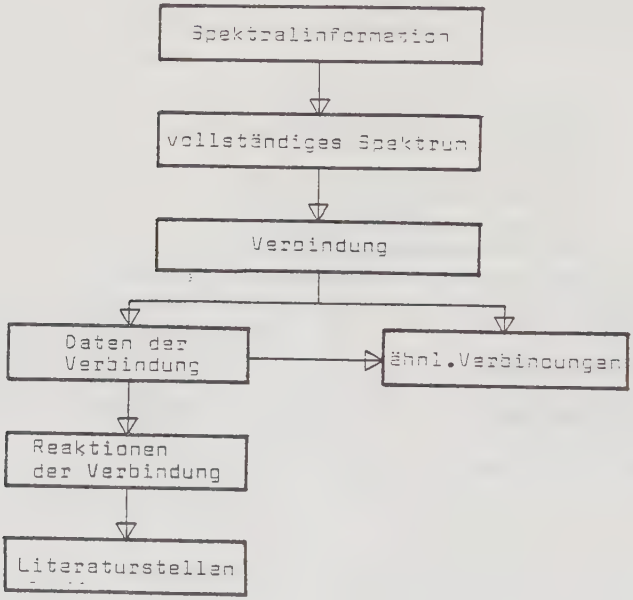
Vorhandene Information: Nummer einer Verbindung



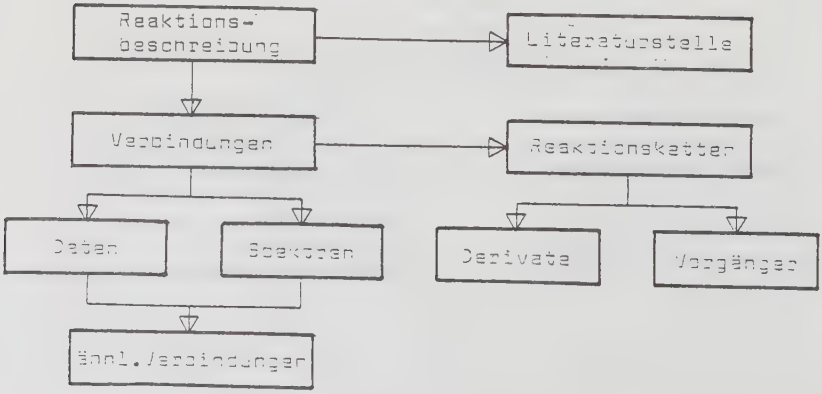
Vorhandene Information: Daten einer Verbindung



Vorhandene Information: Bande bzw. Peak eines IR- bzw. NMR-Spektrums



Vorhandene Information: Reaktionsbeschreibung



3.2.1. Eigenschaften der Verbindungen

Alle unter der Nummer einer chemischen Verbindung gespeicherten Daten lassen mit dem Code

EI,V_a

aufgerufen.

Beispiele:

ei,v 2232
2232: 3-BROM-ZIMTSAEURE
C 9, H 7, O 2, Br
?

ei,v 2280
2280: 6-CHLOR-2-MORPHOLINO-4-(N-PHENYL-AMINO)CHINAZOLIN
C 18, H 17, N 4, O, Cl
SCHMELZPUNKT: 264 - 265 GRAD C

ei,v 2253
2253: 4-METHOXYANILIN
C 7, H 9, N, O
SCHMELZPUNKT: 57 GRAD C
SIEDEPUNKT: 246 GRAD C BEI 1 ATM
pKb: 5,34

ei,v 2202

2202: MORPHOLIN-4-CARBONSAEURE-(4)-(2-METHYLANILID)-(2-MORPHOLINO-8-METHINAZOLYL-(4)-IMID)

C 25, H 30, N 6, O 2

SCHMELZPUNKT: 119 GRAD C

I R - S P E K T R U M (PERKIN-ELMER 337):

MEDIUM: KALIUMBROMID

WELLENZAHL	INTENSITAET	ZUORDNUNG
3120	SCHWACH	N-H-VALENZSCHWINGUNG sekundaeres AMIN
3060	SCHWACH	C-H-VALENZSCHWINGUNG AROMAT
2860	SCHWACH	C-H-VALENZSCHWINGUNG ALIPHAT
1620	SEHR STARK	C-N-DOPPELBINDUNGSVALENZSCHWINGUNG HETEROCYCLUS
1450	SEHR STARK	--
1130	SEHR STARK	C-O-EINFACHBINDUNGSVALENZSCHWINGUNG HETEROCYCLUS
980	STARK	--
765	SCHWACH	C-H-WAGGINGSCHWINGUNG 3 H
735	SCHWACH	C-H-WAGGINGSCHWINGUNG 4 H

1 . N M R - S P E K T R U M (60 MHz):

STANDARD: TETRAMETHYLSILAN

TEMPERATUR: 30 GRAD C

LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
11,2	1	SINGULETT	-	sekundaeres AMIN
8,1	1	DUBLETT	-	AROMAT
6,8 7,6	6	MULTIPLETT	-	AROMAT
3,7	8	SINGULETT	-	HETEROCYCLUS
3,5	8	MULTIPLETT	-	HETEROCYCLUS
2,5	3	SINGULETT	-	METHYLGRUPPE
2,2	3	SINGULETT	-	METHYLGRUPPE

2 . N M R - S P E K T R U M (60 MHz):

STANDARD: TETRAMETHYLSILAN

TEMPERATUR: 30 GRAD C

LOESUNGSM.: DEUTEROCHLOROFORM
DEUTERIUMOXID

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
8,1	1	DUBLETT	-	AROMAT
6,8 7,6	6	MULTIPLETT	-	AROMAT
3,7	8	SINGULETT	-	HETEROCYCLUS
3,4 3,8	8	MULTIPLETT	-	HETEROCYCLUS
2,5	3	SINGULETT	-	METHYLGRUPPE
2,2	3	SINGULETT	-	METHYLGRUPPE

Nach der Summenformel einer Verbindung fragt man mit dem Code

SF,V_a

Beispiel:

2121
1: AZETIDIN-1-CARBONSAEURE-(3.4-DIOXO-2-PHENYL-CYCLOBUTEN-1-ENYLIMID)-CHLORID
C 14, H 11, N 2, O 2, Cl

sf,v 1135
1135: AETHANOL
C 2, H 6, O

Aussagefähiger sind die Antworten, die man erhält, wenn man nach allen Verbindungen sucht, die einer bestimmten, vorgegebenen Summenformel genügen:

V,SF_a

Dabei ist bei der Eingabe der Summenformel zu beachten, daß auch die Atomanzahl "1" eingegeben werden muß:

also nicht: C_2H_6O

sondern : $C_2H_6O_1$

Die erste Summenformel wird durch das System zurückgewiesen.

Beispiel:

```
v,sf c2h6o1
1135:  AETHANOL
       C 2, H 6, O
1125:  DIMETHYLAETHER
       C 2, H 6, O
```

Nach Teilen einer Summenformel und nach Fragmenten einer Formel fragt man mit den Codes

V,TSF_a

und

V,FSF_a

Unter einem Teil einer Summenformel wird verstanden, daß dieser Teil einer Formel mindestens im Molekül enthalten sein muß, man kann also bestimmte Atomsorten und Atomzahlen für eine Suche vorgeben.
Eine Teilsummenformel einer Verbindung wird dann benötigt, wenn man sicher sein möchte, daß eine gewisse Atomanzahl nicht unterschritten wird.

```
v,tsf c6h6ol
 1198:  ACETALDEHYDDIAETHYLACETAL
        C 6, H 14, O 2
 1453:  ACETALDEHYDDIISOPROPYLACETAL
        C 8, H 18, O 2
 1488:  ACETALDEHYDDIPROPYLACETAL
        C 8, H 18, O 2
 1240:  ACETESSIGESTER
        C 6, H 10, O 3
 1287:  ACETONDIAETHYLKETAL
        C 7, H 16, O 2
 1537:  ACETONDIPROPYLKETAL
        C 9, H 20, O 2
 1536:  ACETONMETHYLAETHYLENKETAL
        C 6, H 12, O 2
 1335:  C-ACETYLACETESSIGESTER
        C 8, H 12, O 4
 1694:  ACETYL-AETHYLMALONBENZYLESTER
        C 21, H 22, O 5
 1695:  ACETYL-AETHYLMALONSAEURE
        C 7, H 10, O 5
 1365:  ACETYLBERNSTEINSAEUREAETHYLESTER
        C 10, H 16, O 5
 1561:  ACETYLMALONESTER
        C 9, H 14, O 5
 2082:  3-ACETYL-2-MORPHOLINO-5-PHENYL-3H-PYRIMIDIN-4-ON
        C 16, H 17, N 3, O 3
```

...

Fortsetzung der Antwortliste "v,tsf C6H601" :

```
 1492:  ACROLEINDIAETHYLACETAL
        C 7, H 14, O 2
 1493:  ACROLEINDIPROPYLACETAL
        C 9, H 18, O 2
 1517:  ACROLEINMETHYLAETHYLENACETAL
        C 6, H 10, O 2
 1510:  ACRYLSAEUREANHYDRID
        C 6, H 6, O 3
 1513:  ACRYLSAEUREISOPROPYLESTER
        C 6, H 10, O 2
 1509:  ACRYLSAEURE-PROPIONSAEUREANHYDRID
        C 6, H 8, O 3
 1569:  ACRYLSAEUREPROPYLESTER
        C 6, H 10, O 2
 1361:  AETHYLACETESSIGESTER
        C 8, H 14, O 3
 1396:  AETHYLCHLORMALONSAEUREESTER
        C 9, H 15, O 4, Cl
AUSGABE FORTSETZEN?  (J;N)
j
 1344:  AETHYLIDENDIACETAT
        C 6, H 10, O 4
 1692:  AETHYL-MALONBENZYLESTER
        C 19, H 20, O 4
 1329:  AETHYLMALONESTER
        C 9, H 16, O 4
 1190:  3-AETHYL-PENTANOL-3
        C 7, H 16, O
 1654:  N-AETHYL-PHTHALIMID
        C 10, H 9, N, O 2
 1686:  ALANIN-BENZYLESTER
        C 10, H 13, N, O 2
 1756:  ALANIN-tert.BUTYLESTER
        C 7, H 15, N, O 2
 1584:  ALANYLALANIN
        C 6, H 12, N 2, O 3
 1626:  ALANYLGLYCINAETHYLESTER
        C 7, H 14, N 2, O 3
 1586:  ALANYLLEUCIN
        C 9, H 18, N 2, O 3
```

Den Schmelzpunkt einer bekannten Verbindung erfragt man durch Eingabe des Codes

FP,V_a

Beispiel:

```
2117
7:  PYRROLIDIN-1-CARBONSAEURE-(3.4-DIOXO-2-PHENYL-CYCLOBUT-1-ENYLIMID)-IOD
    ID
    SCHMELZPUNKT:   149 GRAD C
```

Sucht man zu einem bestimmten, durch Messung vorgegebenen Schmelzpunkt alle im System gespeicherten Verbindungen, so erfolgt dies mittels

V,Fp_a

Beispiel:

```
v,fp 100

  1836:  N.N'-DIISOPROPYL-CHLORFORMAMIDIN-HYDROCHLORID
         SCHMELZPUNKT:   97 - 105 GRAD C
```

Alle Temperaturen im System (Schmelzpunkt, Siedepunkt und Reaktionstemperatur) werden in °C nachgewiesen.

Die Temperaturintervalle werden durch den Benutzer bei der Abfrage definiert:

V,Fp_a_1; a_2

Beispiele:

v,fp 212;213

2298: 1-(4-CHLORPHENYL)-2-IMINO-6-MORPHOLINO-4-(TRICHLORMETHYL)-1.2-DIHY
1.3.5-TRIAZIN
SCHMELZPUNKT: 212 GRAD C Z

2290: 2.2.2-TRICHLOR-N-(2-MORPHOLINO-BENO(b)CHINAZOLIN-4-YL)ACETAMIDIN
SCHMELZPUNKT: 213 GRAD C Z

Im folgenden Beispiel ist der erlaubte Temperaturbereich für den Schmelzpunkt größer:

210;220

8: 1-(4-CHLORPHENYL)-2-IMINO-6-MORPHOLINO-4-(TRICHLORMETHYL)-1.2-DIHYDRO-1.3.5-TRIAZIN
SCHMELZPUNKT: 212 GRAD C Z

8: 2.4-DIMORPHOLINO-5.6-DIPHENYL-PYRIMIDIN-HYDROCHLORID
SCHMELZPUNKT: 220 - 222 GRAD C

2: (4-ETHOXY-1-NAPHTHYL)-CARBAMIDSAEUREMORPHOLID
SCHMELZPUNKT: 219 GRAD C

0: 2-IMINO-6-MORPHOLINO-4-PHENYL-1-(p-PHENYLEN)-1.2-DIHYDRO-1.3.5-TRIAZIN
SCHMELZPUNKT: 214 GRAD C

1: 2-MORPHOLINO-3H-THIENO-(2.3-d)-PYRIMIDIN-4-SELON
SCHMELZPUNKT: 210,2 GRAD C

2: 2-MORPHOLINO-3H-THIENO-(2.3-d)-PYRIMIDIN-4-THION
SCHMELZPUNKT: 217 GRAD C

9: N-(2-MORPHOLINO-6-PHENYL-4-CHINAZOLINYL)BENZAMIDIN
SCHMELZPUNKT: 218 GRAD C

0: N-(2-MORPHOLINO-THIENO-(2.3-d)-PYRIMIDIN-4-YL)-N'-(2)-THIENYL-MORPHOLIN-4-CARBAMIDINIUMCHLORID
SCHMELZPUNKT: 215,7 GRAD C

7: alpha-PHENYLIMINO-alpha-MORPHOLINO-ACETAMID
SCHMELZPUNKT: 219 GRAD C

3: N-PHENYL-N'-(2-PHENYL-3.4-DIOXO-CYCLOBUT-1-ENYL)-PYRROLIDIN-1-CARBAMIDIN
SCHMELZPUNKT: 220 GRAD C

2: PYRROLIDIN-1-CARBONSAEURE-2-THIENYLAMID
SCHMELZPUNKT: 214 GRAD C

0: 2.2.2-TRICHLOR-N-(2-MORPHOLINO-BENZO)CHINAZOLIN-4-YL)ACETAMIDIN
SCHMELZPUNKT: 213 GRAD C Z

7: 2.2.2-TRICHLOR-N-(2-MORPHOLINO-4-CHINAZOLINYL)ACETAMIDIN-HYDROCHLORID
SCHMELZPUNKT: 210 GRAD C Z

Für den Siedepunkt einer Verbindung gilt das oben Gesagte entsprechend.

Will man gleichzeitig nach dem Schmelzpukt und nach dem Siedepunkt einer Verbindung fragen, so benutzt man diesen Code:

FP,KP,V_a

Beispiel:

```
fp,kp,v 2119
  2119:  p-ANISIDIN
         SCHMELZPUNKT:    57 GRAD C
         SIEDEPUNKT:      244 GRAD C
```

6.2.2. Spektraldaten

Mit den Codes

SE

und INT

fragt man nach den Listen der Schwingungselemente und der Interpretationen. (Die Listen finden sich bei der Diskussion der Spektrendarstellung)
Ob unter einer Verbindung auch Spektren gespeichert sind, kann man mit

IR,V_a

und NMR,V_a

abfragen.

Beispiele:

2114
: PYRROLIDIN-1-CARBONSAEURE-(3.4-DIOXO-2-PHENYL-CYCLOBUT-1-ENYLIMID)-FLU
ORID
I R - S P E K T R U M (PERKIN-ELMER 177):
MEDIUM: KALIUMBROMID

WELLENZAHL	INTENSITAET	ZUORDNUNG
3065	MITTEL	C-H-VALENZSCHWINGUNG AROMAT
2990	SCHWACH	C-H-VALENZSCHWINGUNG ALIPHAT
2895	SCHWACH	C-H-VALENZSCHWINGUNG ALIPHAT
1790	MITTEL	C-N-DOPPELBINDUNGSVALENZSCHWINGUNG CHLORFORMAMIDIN
1740	STARK	C-O-DOPPELBINDUNGSVALENZSCHWINGUNG 4-RINGKETON
1645	SEHR STARK	--
1120	STARK	--
1060	STARK	--
772	MITTEL	C-H-WAGGINGSCHWINGUNG 5 H
698	MITTEL	C-H-WAGGINGSCHWINGUNG 5 H

Ein ^{1}H-NMR-Spektrum einer bestimmten Verbindung wird so auf dem Bildschirm bzw. auf dem Ausdruck dargestellt:

```
nmr,v 2105
 2105:  1-PHENYL-CYCLOBUTEN-3.4-DION
        N M R - S P E K T R U M  (60 MHz):
        STANDARD:      TETRAMETHYLSILAN
        TEMPERATUR:    30 GRAD C
        LOESUNGSM.:    DEUTEROCHLOROFORM
            ppm        H   STRUKTUR   J (Hz)   ZUORDNUNG
            9,5        1   SINGULETT     -     ALKEN
          7,8   7,9    2   MULTIPLETT    -     AROMAT
          7,3   7,6    3   MULTIPLETT    -     AROMAT
```

Die umgekehrte, sicher häufigere Fragestellung ist die nach Verbindungen, wenn vollständige oder Teilinformationen über die Spektren vorliegen.
Liegt nur eine charakteristische Wellenzahl eines Infrarotspektrums vor, so lautet der Code:

V,WZ_a

Beispiel:

z 1695

57: 3-(N'-METHYL-N'-PHENYL-UREIDO)-THIOPHEN-2-CARBONSAEURE-METHYLESTER
I R - S P E K T R U M (PERKIN-ELMER 177):
MEDIUM: KALIUMBROMID

WELLENZAHL	INTENSITAET	ZUORDNUNG
3310	STARK	N-H-VALENZSCHWINGUNG HARNSTOFF
1695	STARK	C-O-DOPPELBINDUNGSVALENZSCHWINGUNG HARNSTOFF
1675	STARK	C-O-DOPPELBINDUNGSVALENZSCHWINGUNG ESTER

z 3010

09: 2-MORPHOLINO-CHINAZOLON
I R - S P E K T R U M (PERKIN-ELMER 337):
MEDIUM: KALIUMBROMID

WELLENZAHL	INTENSITAET	ZUORDNUNG
3050	MITTEL	N-H-VALENZSCHWINGUNG HETEROCYCLUS
3010	MITTEL	C-H-VALENZSCHWINGUNG AROMAT
2950	MITTEL	C-H-VALENZSCHWINGUNG ALIPHAT
1680	SEHR STARK	C-O-DOPPELBINDUNGSVALENZSCHWINGUNG HETEROCYCLUS
1600	SEHR STARK	C-N-DOPPELBINDUNGSVALENZSCHWINGUNG HETEROCYCLUS
1270	STARK	--
1120	SEHR STARK	C-O-EINFACHBINDUNGSVALENZSCHWINGUNG HETEROCYCLUS
760	MITTEL	C-H-WAGGINGSCHWINGUNG 4 H

Ist man sich über die genaue Lage der Wellenzahlen nicht so sicher oder hat man bei einer Suche nach präzisen Wellenzahlen keine Verbindung gefunden, so besteht die Möglichkeit, den Wellenzahlenbereich zu erweitern:

$$V,wZ_a_1 \ , \ a_2$$

Beispiel:

v,wz 1690,1700

2157: 3-(N'-METHYL-N'-PHENYL-UREIDO)-THIOPHEN-2-CARBONSAEURE-METHYLESTE
I R - S P E K T R U M (PERKIN-ELMER 177):
MEDIUM: KALIUMBROMID

WELLENZAHL	INTENSITAET	ZUORDNUNG
3310	STARK	N-H-VALENZSCHWINGUNG HARNSTOFF
1695	STARK	C-O-DOPPELBINDUNGSVALENZSCHWINGUNG HARNSWOFF
1675	STARK	C-O-DOPPELBINDUNGSVALENZSCHWINGUNG ESTER

2212: 1-(3-NITRO-PHENYL)-2-MORPHOLINO-5-(3-NITRO-PHENYLIMINO-MORPHOLINO
YL)-1.3.5-TRIAZINDION-(4.6)
I R - S P E K T R U M (PERKIN-ELMER 337):
MEDIUM: KALIUMBROMID

WELLENZAHL	INTENSITAET	ZUORDNUNG
3100	SCHWACH	C-H-VALENZSCHWINGUNG AROMAT
2980	MITTEL	C-H-VALENZSCHWINGUNG ALIPHAT
2920	MITTEL	C-H-VALENZSCHWINGUNG ALIPHAT
1700	SEHR STARK	C-O-DOPPELBINDUNGSVALENZSCHWINGUNG HETEROCYCLUS
1500	STARK	--
1140	STARK	C-O-EINFACHBINDUNGSVALENZSCHWINGUNG HETEROCYCLUS
1120	STARK	C-O-EINFACHBINDUNGSVALENZSCHWINGUNG HETEROCYCLUS
780	MITTEL	C-H-WAGGINGSCHWINGUNG 3 H
690	MITTEL	C-H-WAGGINGSCHWINGUNG 3 H

...

Fortsetzung des Beispiels:

2195: PHENYLIMINO-MORPHOLINO-(2-CARBETHOXYHYDRAZINO)-METHAN
I R - S P E K T R U M (PERKIN-ELMER 177):
MEDIUM: KALIUMBROMID

WELLENZAHL	INTENSITAET	ZUORDNUNG
3310	MITTEL	N-H-VALENZSCHWINGUNG AMIN
3200	MITTEL	N-H-VALENZSCHWINGUNG AMIN
1700	SEHR STARK	N-H-VALENZSCHWINGUNG HYDRAZINDERIVAT
1120	SEHR STARK	C-O-EINFACHBINDUNGSVALENZSCHWINGUNG HETEROCYCLUS
755	STARK	C-H-WAGGINGSCHWINGUNG 5 H
690	STARK	C-H-WAGGINGSCHWINGUNG 5 H

Bis maximal zehn einzelne Wellenzahlen kann der Anfrager vorgeben:

$$V,WZN_a_1, \ldots a_{10}$$

Beispiel:

```
v,wzn 3450,3080,2820,1615,1515,1325,990,870,757,695
```

1880: N-METHYLANILIN
I R - S P E K T R U M (ALDRICH-KATALOG 1975):
MEDIUM: NATRIUMCHLORID

WELLENZAHL	INTENSITAET	ZUORDNUNG
3450	SEHR STARK	N-H-VALENZSCHWINGUNG sekundaeres AMIN
3080	MITTEL	C-H-VALENZSCHWINGUNG AROMAT
2820	MITTEL	C-H-VALENZSCHWINGUNG METHYLGRUPPE
1615	MITTEL	C-C-DOPPELBINDUNGSVALENZSCHWINGUNG AROMAT
1515	MITTEL	--
1325	MITTEL	--
990	MITTEL	--
870	MITTEL	--
757	MITTEL	C-H-WAGGINGSCHWINGUNG 5 H
695	MITTEL	C-H-WAGGINGSCHWINGUNG 5 H

Der Fall, daß von einer Verbindung nur eine oder mehrere Spektralbanden oder weitere physikalische Daten bekannt sind, ist in der Laborpraxis selten. In vielen Fällen ist zumindest ein Teil der Struktur bekannt,da man im allgemeinen weiß,welche Edukte reagiert haben oder welche Reaktion stattgefunden hat, so daß man zumindest über einige Teilinformationen verfügt.
Man kann also nach dem gesuchten Schwingungselement fragen:

V,SE_a

Das gesuchte Schwingungselement wird mit der Codezahl aus der Schwingungselementliste verschlüsselt.

Beispiel:

v,se 100

2082: 3-ACETYL-2-MORPHOLINO-5-PHENYL-3H-PYRIMIDIN-4-ON
I R - S P E K T R U M (PERKIN-ELMER 177):
MEDIUM: KALIUMBROMID

WELLENZAHL	INTENSITAET	ZUORDNUNG
2900	MITTEL	C-H-VALENZSCHWINGUNG ALIPHAT
1760	STARK	C-O-DOPPELBINDUNGSVALENZSCHWINGUNG KETON
1610	STARK	C-O-DOPPELBINDUNGSVALENZSCHWINGUNG 6-RINGKETON
1150	STARK	C-O-EINFACHBINDUNGSVALENZSCHWINGUNG cyclischer ETHER
690	MITTEL	C-H-WAGGINGSCHWINGUNG 5 H

...

(Schwingungselement 100 = Aliphat)

Fortsetzung des Beispiels "V,SE 1oo"

3-CHLOR-N- (DIMETHYLAMINO) (4-METHOXYPHENYLAMINO)METHYLEN)-6-METHYLBENZO(b)THIOPHEN-2-CARBOXAMID

I R - S P E K T R U M (PERKIN-ELMER 177):

MEDIUM: KALIUMBROMID

WELLENZAHL	INTENSITAET	ZUORDNUNG
3220	MITTEL	N-H-VALENZSCHWINGUNG sekundaeres AMIN
3085	MITTEL	C-H-VALENZSCHWINGUNG AROMAT
2940	MITTEL	C-H-VALENZSCHWINGUNG ALIPHAT
2840	MITTEL	C-H-VALENZSCHWINGUNG METHYLGRUPPE
1610	STARK	C-O-DOPPELBINDUNGSVALENZSCHWINGUNG CARBONYLVERBINDUNGEN
780	MITTEL	C-H-WAGGINGSCHWINGUNG 3 H
740	MITTEL	C-H-WAGGINGSCHWINGUNG 3 H

3-CHLOR-N-((DIMETHYLAMINO)(p-TOLYLTHIO)METHYLEN)-6-METHYL-BENZO(b)THIOPHEN-2-CARBOXAMID

I R - S P E K T R U M (PERKIN-ELMER 177):

MEDIUM: KALIUMBROMID

WELLENZAHL	INTENSITAET	ZUORDNUNG
2920	SCHWACH	C-H-VALENZSCHWINGUNG METHYLGRUPPE
1600	STARK	C-O-DOPPELBINDUNGSVALENZSCHWINGUNG CARBONYLVERBINDUNGEN
1580	STARK	C-N-DOPPELBINDUNGSVALENZSCHWINGUNG AMIDIN
760	MITTEL	C-H-WAGGINGSCHWINGUNG 3 H
720	MITTEL	C-H-WAGGINGSCHWINGUNG 3 H

3-CHLOR-N-DIMETHYLCARBAMOYL-6-METHYLBENZO(b)THIOPHEN-2-CARBOXAMID

I R - S P E K T R U M (PERKIN-ELMER 177):

MEDIUM: KALIUMBROMID

WELLENZAHL	INTENSITAET	ZUORDNUNG
3300	STARK	N-H-VALENZSCHWINGUNG HARNSTOFF
2940	STARK	C-H-VALENZSCHWINGUNG ALIPHAT
1685	SEHR STARK	C-O-DOPPELBINDUNGSVALENZSCHWINGUNG HARNSTOFF
1670	STARK	C-O-DOPPELBINDUNGSVALENZSCHWINGUNG HARNSTOFF
1595	STARK	C-C-DOPPELBINDUNGSVALENZSCHWINGUNG AROMAT
760	MITTEL	C-H-WAGGINGSCHWINGUNG AROMAT
740	MITTEL	C-H-WAGGINGSCHWINGUNG AROMAT

Fragt man genauer nach den Interpretationen, so wird im Abfragecode spezifiziert, ob das IR- oder NMR-Spektrum angesprochen werden soll:

V,IR,INT_a

oder

V,NMR,INT_a

Beispiel:

(Interpretation 23o = Keton)

v,ir,int 230

2082: 3-ACETYL-2-MORPHOLINO-5-PHENYL-3H-PYRIMIDIN-4-ON
I R - S P E K T R U M (PERKIN-ELMER 177):
MEDIUM: KALIUMBROMID

WELLENZAHL	INTENSITAET	ZUORDNUNG
2900	MITTEL	C-H-VALENZSCHWINGUNG ALIPHAT
1760	STARK	C-O-DOPPELBINDUNGSVALENZSCHWINGUNG KETON
1610	STARK	C-O-DOPPELBINDUNGSVALENZSCHWINGUNG 6-RINGKETON
1150	STARK	C-O-EINFACHBINDUNGSVALENZSCHWINGUNG cyclischer ETHER
690	MITTEL	C-H-WAGGINGSCHWINGUNG 5 H

2211: 2-MORPHOLINO-6-CHLOR-CHINAZOLON-(4)
I R - S P E K T R U M (PERKIN-ELMER 337):
MEDIUM: KALIUMBROMID

WELLENZAHL	INTENSITAET	ZUORDNUNG
3250	SCHWACH	N-H-VALENZSCHWINGUNG AMIN
3050	SCHWACH	C-H-VALENZSCHWINGUNG AROMAT
2950	SCHWACH	C-H-VALENZSCHWINGUNG ALIPHAT
1680	SEHR STARK	C-O-DOPPELBINDUNGSVALENZSCHWINGUNG KETON
500	SEHR STARK	C-N-DOPPELBINDUNGSVALENZSCHWINGUNG HETEROCYCLUS
1120	STARK	C-O-EINFACHBINDUNGSVALENZSCHWINGUNG HETEROCYCLUS
825	MITTEL	--
650	SCHWACH	C-H-WAGGINGSCHWINGUNG 3 H
570	SCHWACH	C-H-WAGGINGSCHWINGUNG 3 H

Die kombinierte Abfrage nach einer Interpretation und einer Wellenzahl erlaubt es, die vorhandene Strukturinformation mit der Information aus der Messung zu kombinieren:

V,INT,WZ_a, b

Die Angabe "IR" wird weggelassen, da es selbstverständlich ist, daß sich die Abfrage auf ein IR-Spektrum bezieht, wenn die zusätzliche Spezifikation "Wellenzahl" vorliegt.

Beispiel:

```
wz 473,1640

3-CHLOR-N-(CHLOR(DIMETHYLAMINO)-METHYLEN)-6-METHYL-BENZO(b)THIOPHEN-2-
CARBOXAMID
I R - S P E K T R U M   (PERKIN-ELMER 177):
MEDIUM:        KALIUMBROMID
WELLENZAHL    INTENSITAET    ZUORDNUNG
  3020         SCHWACH        C-H-VALENZSCHWINGUNG
                              AROMAT
  2985         SCHWACH        C-H-VALENZSCHWINGUNG
                              METHYLGRUPPE
  1640         STARK          C-O-DOPPELBINDUNGSVALENZSCHWINGUNG
                              CHLORFORMAMIDIN
  1585         STARK          C-N-DOPPELBINDUNGSVALENZSCHWINGUNG
                              CHLORFORMAMIDIN

3-CHLOR-N-(CHLOR(1-PYRROLIDINYL)-METHYLEN)-6-METHYL-BENZO(b)THIOPHEN-2
-CARBOXAMID
I R - S P E K T R U M   (PERKIN-ELMER 177):
MEDIUM:        KALIUMBROMID
WELLENZAHL    INTENSITAET    ZUORDNUNG
  2980         SCHWACH        C-H-VALENZSCHWINGUNG
                              ALIPHAT
  2895         SCHWACH        C-H-VALENZSCHWINGUNG
                              ALIPHAT
  1640         STARK          C-O-DOPPELBINDUNGSVALENZSCHWINGUNG
                              CHLORFORMAMIDIN
  1585         STARK          C-N-DOPPELBINDUNGSVALENZSCHWINGUNG
                              CHLORFORMAMIDIN
                                                          ...
```

Fortsetzung des Beispiels "V,INT,WZ 473,164a":

2075: 1.5-bis-(CHLOR-MORPHOLINO-METHYLENAMINO)-NAPHTHALIN
I R - S P E K T R U M (PERKIN-ELMER 177):

MEDIUM:	KALIUMBROMID	
WELLENZAHL	INTENSITAET	ZUORDNUNG
1640	SEHR STARK	C-N-DOPPELBINDUNGSVALENZSCHWINGUNG CHLORFORMAMIDIN

2086: 4-(N-MESITYL-CHLORFORMIMIDOYL)-MORPHOLIN
I R - S P E K T R U M (PERKIN-ELMER 177):

MEDIUM:	KALIUMBROMID	
WELLENZAHL	INTENSITAET	ZUORDNUNG
1640	STARK	C-N-DOPPELBINDUNGSVALENZSCHWINGUNG CHLORFORMAMIDIN

2055: (6-METHYL-3-CHLOR-2-CARBONYLBENZO-(b)-THIOPHENO)-MORPHOLIN-4-CARBIM
YLCHLORID
I R - S P E K T R U M (PERKIN-ELMER 177):

MEDIUM:	KALIUMBROMID	
WELLENZAHL	INTENSITAET	ZUORDNUNG
1640	MITTEL	C-O-DOPPELBINDUNGSVALENZSCHWINGUNG CHLORFORMAMIDIN
1595	MITTEL	C-N-DOPPELBINDUNGSVALENZSCHWINGUNG CHLORFORMAMIDIN

2083: 4-(N-(1-NAPHTHYL)-CHLORFORMIMIDOYL)-MORPHOLIN
I R - S P E K T R U M (PERKIN-ELMER 177):

MEDIUM:	KALIUMBROMID	
WELLENZAHL	INTENSITAET	ZUORDNUNG
1640	MITTEL	C-N-DOPPELBINDUNGSVALENZSCHWINGUNG CHLORFORMAMIDIN

Es kann auch bei dieser Abfrage sinnvoll sein, den Wellenzahlbereich zu erweitern:

$$V,INT,WZ_a,b_1,b_2$$

$b_1 - b_2$ beschreiben den Wellenzahlbereich, in dem gesucht wird.

Beispiel:

```
:,wz 473,1630,1650
```

5: AZETIDIN-1-CARBONSAEURE-(3.4-DIOXO-2-CHLOR-CYCLOBUT-1-ENYLIMID)-CHLORID

I R - S P E K T R U M (PERKIN-ELMER 177):

MEDIUM:	KALIUMBROMID	
WELLENZAHL	INTENSITAET	ZUORDNUNG
3005	SCHWACH	C-H-VALENZSCHWINGUNG ALIPHAT
2940	SCHWACH	C-H-VALENZSCHWINGUNG ALIPHAT
1795	STARK	C-O-DOPPELBINDUNGSVALENZSCHWINGUNG 4-RINGKETON
1740	STARK	C-O-DOPPELBINDUNGSVALENZSCHWINGUNG 4-RINGKETON
1630	STARK	C-N-DOPPELBINDUNGSVALENZSCHWINGUNG CHLORFORMAMIDIN
1540	MITTEL	C-N-DOPPELBINDUNGSVALENZSCHWINGUNG CHLORFORMAMIDIN
1280	MITTEL	--
1140	MITTEL	--
618	MITTEL	C-Halogen-VALENZSCHWINGUNG CHLORID

4: 4-(N-(4-BIPHENYLYL)-CHLORFORMIMIDOYL)-MORPHOLIN

I R - S P E K T R U M (PERKIN-ELMER 177):

MEDIUM:	KALIUMBROMID	
WELLENZAHL	INTENSITAET	ZUORDNUNG
1650	MITTEL	C-N-DOPPELBINDUNGSVALENZSCHWINGUNG CHLORFORMAMIDIN

...

Fortsetzung des Beispiels "V,INT,WZ 473,163o,165o":

2063: 6-BROM-3-CHLOR-N-(CHLOR(DIMETHYLAMINO)-METHYLEN)-BENZO(b)THIOPHEN-RBOXAMID

I R - S P E K T R U M (PERKIN-ELMER 177):

MEDIUM: KALIUMBROMID

WELLENZAHL	INTENSITAET	ZUORDNUNG
3035	SCHWACH	C-H-VALENZSCHWINGUNG AROMAT
2975	SCHWACH	C-H-VALENZSCHWINGUNG ALIPHAT
1635	STARK	C-O-DOPPELBINDUNGSVALENZSCHWINGUNG CHLORFORMAMIDIN
1545	STARK	C-N-DOPPELBINDUNGSVALENZSCHWINGUNG CHLORFORMAMIDIN

1896: 3-CHLOR-N-(CHLOR(DIMETHYLAMINO)-METHYLEN)-BENZO(b)THIOPHEN-2-CARBOD

I R - S P E K T R U M (PERKIN-ELMER 177):

MEDIUM: KALIUMBROMID

WELLENZAHL	INTENSITAET	ZUORDNUNG
3025	SCHWACH	C-H-VALENZSCHWINGUNG AROMAT
2985	SCHWACH	C-H-VALENZSCHWINGUNG ALIPHAT
1645	STARK	C-O-DOPPELBINDUNGSVALENZSCHWINGUNG CHLORFORMAMIDIN
1615	STARK	C-N-DOPPELBINDUNGSVALENZSCHWINGUNG CHLORFORMAMIDIN

2002: 3-CHLOR-N-(CHLOR(DIMETHYLAMINO)-METHYLEN)-6-CYAN-BENZO(b)THIOPHEN-RBOXAMID

I R - S P E K T R U M (PERKIN-ELMER 177):

MEDIUM: KALIUMBROMID

WELLENZAHL	INTENSITAET	ZUORDNUNG
3025	SCHWACH	C-H-VALENZSCHWINGUNG AROMAT
2215	SEHR STARK	C-N-DREIFACHBINDUNGSVALENZSCHWINGUNG NITRIL
1650	STARK	C-O-DOPPELBINDUNGSVALENZSCHWINGUNG CHLORFORMAMIDIN
1600	STARK	C-N-DOPPELBINDUNGSVALENZSCHWINGUNG CHLORFORMAMIDIN

Je mehr Wellenzahlen man explizit vorgibt, desto größer wird die Wahrscheinlichkeit, kein Ergebnis zu erhalten.

Beispiel:

```
v,wzn 3440,3090,2800,1620,1520,1330,980,860,760,690
NICHT IM BESTAND
```

Um nun nachzuprüfen, ob dieses Ergebnis durch den eventuell zu geringen Datenbestand oder durch eine chemisch nicht mehr sinnvolle Häufung von Beschreibungselementen zustande kommt, kann man den Fragebereich weniger spezifisch fassen, wobei man Vorinformationen aufgibt oder man kann die Beschreibungselemente um einen Mittelwert schwanken lassen.

Code:

$$V,WZN_a_1, \ldots a_{10}(b)$$

Beispiel:

```
,wzn 3440,3090,2800,1620,1520,1330,980,860,760,690 (20)
```

1880: N-METHYLANILIN

I R - S P E K T R U M (ALDRICH-KATALOG 1975):

MEDIUM: NATRIUMCHLORID

WELLENZAHL	INTENSITAET	ZUORDNUNG
3450	SEHR STARK	N-H-VALENZSCHWINGUNG sekundaeres AMIN
3080	MITTEL	C-H-VALENZSCHWINGUNG AROMAT
2820	MITTEL	C-H-VALENZSCHWINGUNG METHYLGRUPPE
1615	MITTEL	C-C-DOPPELBINDUNGSVALENZSCHWINGUNG AROMAT
1515	MITTEL	--
1325	MITTEL	--
990	MITTEL	--
870	MITTEL	--
757	MITTEL	C-H-WAGGINGSCHWINGUNG 5 H
695	MITTEL	C-H-WAGGINGSCHWINGUNG 5 H

Auch mit den Schwingungselementen, den etwas unspezifischeren Strukturbeschreibungsmerkmalen, kann in Kombination mit einem Wellenzahlbereich abgefragt werden:

V,SE,WZ_a,b_1,b_2

Beispiel:

```
v,se,wz 120,3500,3000
```

2169: N-(AMINO(DIMETHYLAMINO)METHYLEN)-3-CHLOR-6-METHYLBENZO(b)THIOPHEN-RBOXAMID

I R - S P E K T R U M (PERKIN-ELMER 177):

MEDIUM: KALIUMBROMID

WELLENZAHL	INTENSITAET	ZUORDNUNG
3260	MITTEL	N-H-VALENZSCHWINGUNG AMIN
2920	MITTEL	C-H-VALENZSCHWINGUNG ALIPHAT
1620	STARK	C-O-DOPPELBINDUNGSVALENZSCHWINGUNG CARBONYLVERBINDUNGEN
1580	STARK	C-N-DOPPELBINDUNGSVALENZSCHWINGUNG AMID

Die Kombination aus Spektraldaten und physikalischen Eigenschaften läßt die folgende Abfrage zu:

$$V,IR,FP,INT_a_1;a_2;b_1;\ \ldots b_5$$

Beispiel:

,fp,int 130;150;120;473

7: 3.4.6-TRICHLOR-N-(CHLOR(DIETHYLAMINO)-METHYLEN)-BENZO(b)THIOPHEN-2-CARBOXAMID
SCHMELZPUNKT: 149,7 GRAD C
I R - S P E K T R U M (PERKIN-ELMER 177):
MEDIUM: KALIUMBROMID

WELLENZAHL	INTENSITAET	ZUORDNUNG
2930	SCHWACH	C-H-VALENZSCHWINGUNG ALKAN
2880	SCHWACH	C-H-VALENZSCHWINGUNG ALKAN
1650	STARK	C-O-DOPPELBINDUNGSVALENZSCHWINGUNG CHLORFORMAMIDIN
1595	STARK	C-N-DOPPELBINDUNGSVALENZSCHWINGUNG CHLORFORMAMIDIN

(Se 12o = Alkan)

(Se 473 = Chlorformamidin)

Die Abfrage nach Daten aus dem NMR-Spektrum ist analog den Abfragen nach Daten aus dem IR-Spektrum aufgebaut:

V,CV_a

(CV = chemische Verschiebung)

Beispiel:

v,cv 4,0

2213: 1-(4-CHLORPHENYL)-2-MORPHOLINO-5-(4-CHLORPHENYLIMINO-MORPHOLINO-ME
)-1.3.5-TRIAZINDION-(4.6)
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
6,7 7,8	8	MULTIPLETT	-	AROMAT
3,3 4,0	16	MULTIPLETT	-	HETEROCYCLUS

2137: 2.4-DIMORPHOLINO-5-PHENYL-PYRIMIDIN
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
7,9	1	SINGULETT	-	ALKEN
7,0 7,5	5	MULTIPLETT	-	AROMAT
4,0	8	SINGULETT	-	HETEROCYCLUS
3,0	8	MULTIPLETT	-	HETEROCYCLUS

1798: 3.3'-DIPHENYL-N.N'-(3.3'-DIMETHOXY-4.4'-DIYL)-bis-PROPIOLIMIDSAEUR
MORPHOLID
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
7,1 7,4	16	MULTIPLETT	-	AROMAT
4,0	6	SINGULETT	-	METHOXYGRUPPE
3,9	16	SINGULETT	-	HETEROCYCLUS

1802: (4-ETHOXY-1-NAPHTHYL)-CARBAMIDSAEUREMORPHOLID
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
6,8 8,2	6	MULTIPLETT	-	AROMAT
6,8	1	SINGULETT	-	HARNSTOFF
3,8 4,2	2	QUARTETT	-	METHYLENGRUPPE
2,8 3,5	8	QUARTETT	-	ALIPHAT
1,4 1,8	3	TRIPLETT	-	METHYLGRUPPE

...

Fortsetzung des Beispiels "V,CV 4,o"

4: N-(4-ETHOXY-1-NAPHTHYL-PHENYL)-3-PHENYL-PROPIOLIMIDSAEURE-MORPHOLID

N M R - S P E K T R U M (60 MHz):

STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
6,8 8,5	11	MULTIPLETT	-	AROMAT
3,9 4,4	2	QUARTETT	-	METHYLENGRUPPE
3,8	8	MULTIPLETT	-	ALIPHAT
1,5 1,9	3	TRIPLETT	-	METHYLGRUPPE

4: N'-(4-ETHOXYPHENYL)-N-METHYL-3.N-DIPHENYLPROPIOLAMIDIN

N M R - S P E K T R U M (60 MHz):

STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
6,8 7,7	14	MULTIPLETT	-	AROMAT
3,8 4,2	2	QUARTETT	-	METHYLENGRUPPE
3,5	3	SINGULETT	-	METHYLGRUPPE
1,2 1,6	3	TRIPLETT	-	METHYLGRUPPE

- N-(4-METHYL-MERCAPTOPHENYL)-3-PHENYL-PROPIOLIMIDSAEURE-MORPHOLID

N M R - S P E K T R U M (60 MHz):

STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
6,8 7,5	9	MULTIPLETT	-	AROMAT
3,7 4,1	8	SINGULETT	-	ALIPHAT
2,5	3	SINGULETT	-	METHYLGRUPPE

0: MORPHOLIN-4-CARBONSAEURE-(3.4-DIOXO-2-PHENYL-CYCLOBUT-1-ENYLIMID)-CHLO
RID

N M R - S P E K T R U M (60 MHz):

STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
7,9 8,3	2	MULTIPLETT	-	AROMAT
7,3 7,6	3	MULTIPLETT	-	AROMAT
3,9 4,1	8	MULTIPLETT	-	HETEROCYCLUS

5: MORPHOLIN-4-CARBONSAEURE-trans-STYRYLAMID

N M R - S P E K T R U M (60 MHz):

STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
7,5	1	DUBLETT	-	ALKEN
6,8 7,5	5	MULTIPLETT	-	AROMAT
6,0	1	DUBLETT	-	ALKEN
3,2 4,0	8	DUBLETT	-	ALIPHAT

Maximal zehn chemische Verschiebungen (ein in CHEMIS gespeichertes Spektrum kann aus 1o Peaks bestehen) sind für die Abfrage vorgebbar:

$$V,CVN_a_1; \ldots a_{1o}$$

Beispiel:
(Es wurden 4 Peaks vorgegeben)

```
v,cvn 7,5;7,3;6,5;3,6
```

```
1784:  N-trans-STYRYL-MORPHOLIN-4-CARBIMIDOYLCHLORID
       N M R - S P E K T R U M  (60 MHz):
       STANDARD:      TETRAMETHYLSILAN
       TEMPERATUR:    30 GRAD C
       LOESUNGSM.:    DEUTEROCHLOROFORM
           ppm        H   STRUKTUR   J (Hz)   ZUORDNUNG
           7,5        1   DUBLETT       -     ALKEN
           7,3        5   MULTIPLETT    -     AROMAT
           6,5        1   DUBLETT       -     ALKEN
           3,6        8   SINGULETT     -     ALIPHAT
```

Die mögliche Schwankungsbreite läßt sich spezifizieren:

$$V,CVN_a_1; a_2; \ldots a_{10} \ (\square)$$

Beispiel:

n 7,5;6,5;3,5(0,3)

5: 4- (4-CHLORPHENYL-IMINO-MORPHOLINO-METHYL)-4-CHLORPHENYL-CARBAMOYL-MOR
PHOLIN)
N M R - S P E K T R U M (60 MHz):
STANDARD: DEUTEROCHLOROFORM
TEMPERATUR: 30 GRAD C
LOESUNGSM.: TETRAMETHYLSILAN

ppm		H	STRUKTUR	J (Hz)	ZUORDNUNG
6,6	7,4	8	MULTIPLETT	-	AROMAT
2,7	3,7	16	MULTIPLETT	-	HETEROCYCLUS

: 1-(4-CHLORPHENYL)-2-MORPHOLINO-5-(4-CHLORPHENYLIMINO-MORPHOLINO-METHYL
)-1.3.5-TRIAZINDION-(4.6)
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm		H	STRUKTUR	J (Hz)	ZUORDNUNG
6,7	7,8	8	MULTIPLETT	-	AROMAT
3,3	4,0	16	MULTIPLETT	-	HETEROCYCLUS

3: 1.2-DIPHENYL-VINYL-MORPHOLIN-4-CARBIMIDOYLCHLORID
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm		H	STRUKTUR	J (Hz)	ZUORDNUNG
6,8	7,8	10	MULTIPLETT	-	AROMAT
6,3		1	SINGULETT	-	ALKEN
3,8		8	SINGULETT	-	ALIPHAT

2: (4-ETHOXY-1-NAPHTHYL)-CARBAMIDSAEUREMORPHOLID
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm		H	STRUKTUR	J (Hz)	ZUORDNUNG
6,8	8,2	6	MULTIPLETT	-	AROMAT
6,8		1	SINGULETT	-	HARNSTOFF
3,8	4,2	2	QUARTETT	-	METHYLENGRUPPE
2,8	3,5	8	QUARTETT	-	ALIPHAT
1,4	1,8	3	TRIPLETT	-	METHYLGRUPPE

...

Fortsetzung des Beispiels "V,CVN 7,5; 6,5; 3,5; (a,3)"

2034: N-(4-ETHOXY-1-NAPHTHYL-PHENYL)-3-PHENYL-PROPIOLIMIDSAEURE-MORPHOLI
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
6,8 8,5	11	MULTIPLETT	-	AROMAT
3,9 4,4	2	QUARTETT	-	METHYLENGRUPPE
3,8	8	MULTIPLETT	-	ALIPHAT
1,5 1,9	3	TRIPLETT	-	METHYLGRUPPE

2044: N'-(4-ETHOXYPHENYL)-N-METHYL-3.N-DIPHENYLPROPIOLAMIDIN
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
6,8 7,7	14	MULTIPLETT	-	AROMAT
3,8 4,2	2	QUARTETT	-	METHYLENGRUPPE
3,5	3	SINGULETT	-	METHYLGRUPPE
1,2 1,6	3	TRIPLETT	-	METHYLGRUPPE

2043: N'-(4-METHOXYPHENYL)-N-METHYL-3.N-DIPHENYLPROPIOLAMIDIN
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
6,8 7,5	14	MULTIPLETT	-	AROMAT
3,6	3	SINGULETT	-	METHOXYGRUPPE
3,4	3	SINGULETT	-	METHYLGRUPPE

2040: N-METHYL-N.N'-DIPHENYL-PROPIOLAMIDIN
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
6,7 7,6	10	MULTIPLETT	-	AROMAT
3,5	3	SINGULETT	-	METHYLGRUPPE
2,7	1	SINGULETT	-	ALKIN

2033: N-(4-METHYL-MERCAPTOPHENYL)-3-PHENYL-PROPIOLIMIDSAEURE-MORPHOLID
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
6,8 7,5	9	MULTIPLETT	-	AROMAT
3,7 4,1	8	SINGULETT	-	ALIPHAT
2,5	3	SINGULETT	-	METHYLGRUPPE

In der Liste "Interpretationen" stehen die Merkmale für IR- und NMR-Spektren nebeneinander, dies wurde eingeführt, da man vermeiden wollte, die Zahl der zu eröffnenden und zu pflegenden Listen zu sehr anwachsen zu lassen. Daher unterscheiden sich die Abfragecodes nur durch die Bezeichnung der Spektren:

V,NMR,INT_a

Beispiel:

v,nmr,int 130

2137: 2.4-DIMORPHOLINO-5-PHENYL-PYRIMIDIN
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
7,9	1	SINGULETT	-	ALKEN
7,0 7,5	5	MULTIPLETT	-	AROMAT
4,0	8	SINGULETT	-	HETEROCYCLUS
3,0	8	MULTIPLETT	-	HETEROCYCLUS

1793: 1.2-DIPHENYL-VINYL-MORPHOLIN-4-CARBIMIDOYLCHLORID
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
6,8 7,8	10	MULTIPLETT	-	AROMAT
6,3	1	SINGULETT	-	ALKEN
3,8	8	SINGULETT	-	ALIPHAT

1794: MORPHOLIN-4-CARBONSAEURE-1.2-DIPHENYL-VINYLAMID
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: HEXADEUTERODIMETHYLSULFOXID

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
8,2	1	SINGULETT	-	HARNSTOFF
7,0 7,8	10	MULTIPLETT	-	AROMAT
6,5	1	SINGULETT	-	ALKEN
3,5	8	MULTIPLETT	-	ALKAN

...

Fortsetzung des Beispiels "V,NMR,INT 13o"

2080: 2-MORPHOLINO-5-PHENYL-3H-PYRIMIDIN-4-ON
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
11,5	1	SINGULETT	-	sekundaeres AMIN
8,1	1	SINGULETT	-	ALKEN
7,1 7,9	5	MULTIPLETT	-	AROMAT
3,8	8	SINGULETT	-	ALIPHAT

2131: 2-MORPHOLINO-5-PHENYL-3H-PYRIMIDIN-4-THION
1 . N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
11,5	1	SINGULETT	-	sekundaeres AMIN
8,0	1	SINGULETT	-	ALKEN
7,0 7,5	5	MULTIPLETT	-	AROMAT
3,7	8	SINGULETT	-	ALIPHAT

2 . N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM
DEUTERIUMOXID

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
8,0	1	SINGULETT	-	ALKEN
7,0 7,5	5	MULTIPLETT	-	AROMAT
3,7	8	SINGULETT	-	ALIPHAT

2105: 1-PHENYL-CYCLOBUTEN-3.4-DION
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
9,5	1	SINGULETT	-	ALKEN
7,8 7,9	2	MULTIPLETT	-	AROMAT
7,3 7,6	3	MULTIPLETT	-	AROMAT

1790: E-PHENYLZIMTSAEURECHLORID
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
8,1	1	SINGULETT	-	ALKEN
6,8 7,6	10	MULTIPLETT	-	AROMAT

(Interpretation 13o = Alken)

r,int 121

9: ACETALDEHYD
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
4,9	1	SINGULETT	-	ALKYLALDEHYD
1,6	3	SINGULETT	-	METHYLGRUPPE

2: 3-ACETYL-2-MORPHOLINO-5-PHENYL-3H-PYRIMIDIN-4-ON
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
8,4	1	SINGULETT	-	AROMAT
7,4	5	SINGULETT	-	AROMAT
3,8	8	SINGULETT	-	ALIPHAT
2,2	3	SINGULETT	-	METHYLGRUPPE

5: AETHANOL
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: AETHANOL

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
3,7	2	QUARTETT	-	METHYLENGRUPPE
2,6	1	SINGULETT	-	primaerer ALKOHOL
1,2	3	TRIPLETT	-	METHYLGRUPPE

9: N-(AMINO(DIMETHYLAMINO)METHYLEN)-3-CHLOR-6-METHYLBENZO(b)THIOPHEN-2-CARBOXAMID
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROTRIFLUORESSIGSAEURE

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
8,0 8,7	2	MULTIPLETT	-	AMIN
7,1 7,7	3	MULTIPLETT	-	AROMAT
3,0	6	SINGULETT	-	METHYLGRUPPE
2,4	3	SINGULETT	-	METHYLGRUPPE

(Int 121 = Methylgruppe)

Die Abfrage nach NMR-Spektraldaten und physikalischen Eigenschaften ist ebenfalls entsprechend den Möglichkeiten strukturiert, die bei den Infrarotspektren bestehen:

V,NMR,FP,INT_a_1; a_2; b_1; ... b_6

und

V,NMR,KP,INT_a_1; a_2; b_1; ... b_6

Beispiele:

v,nmr,fp,int 90;150;110;121

2082: 3-ACETYL-2-MORPHOLINO-5-PHENYL-3H-PYRIMIDIN-4-ON
SCHMELZPUNKT: 110 GRAD C
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
8,4	1	SINGULETT	-	AROMAT
7,4	5	SINGULETT	-	AROMAT
3,8	8	SINGULETT	-	ALIPHAT
2,2	3	SINGULETT	-	METHYLGRUPPE

2172: 3-CHLOR-N.N'-bis(CHLOR(DIMETHYLAMINO)METHYLEN)BENZO b)THIOPHEN-2.6-
ARBOXAMID
SCHMELZPUNKT: 145,2 GRAD C
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROTRIFLUORESSIGSAEURE

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
7,3 7,8	3	MULTIPLETT	-	AROMAT
3,8	12	SINGULETT	-	METHYLGRUPPE

2177: 3-CHLOR-N- (DIMETHYLAMINO)(p-TOLYLTHIO)METHYLEN)-6-METHYL-BENZO b)'
PHEN-2-CARBOXAMID
SCHMELZPUNKT: 128 GRAD C
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROTRIFLUORESSIGSAEURE

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
6,9 7,8	7	MULTIPLETT	-	AROMAT
3,3	6	SINGULETT	-	METHYLGRUPPE
2,5	3	SINGULETT	-	METHYLGRUPPE
2,0	3	SINGULETT	-	METHYLGRUPPE

...

Fortsetzung des Beispiels "V,NMR,FP,INT 9a; 15a; 11a; 121"

: N-(4-METHYLMERCAPTOPHENYL)-CARBAMIDSAEUREMORPHOLID
SCHMELZPUNKT: 126 GRAD C
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
7,3	1	SINGULETT	-	HARNSTOFF
7,0 7,3	4	SINGULETT	-	AROMAT
3,2 3,6	8	DUBLETT	-	ALIPHAT
2,4	3	SINGULETT	-	METHYLGRUPPE

: MORPHOLIN-4-CARBONSAEURE-(4)-(2-METHYLANILID)-(2-MORPHOLINO-8-METHYLCH
INAZOLYL-(4)-IMID)
SCHMELZPUNKT: 119 GRAD C
1 . N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
11,2	1	SINGULETT	-	sekundaeres AMIN
8,1	1	DUBLETT	-	AROMAT
6,8 7,6	6	MULTIPLETT	-	AROMAT
3,7	8	SINGULETT	-	HETEROCYCLUS
3,5	8	MULTIPLETT	-	HETEROCYCLUS
2,5	3	SINGULETT	-	METHYLGRUPPE
2,2	3	SINGULETT	-	METHYLGRUPPE

2 . N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM
DEUTERIUMOXID

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
8,1	1	DUBLETT	-	AROMAT
6,8 7,6	6	MULTIPLETT	-	AROMAT
3,7	8	SINGULETT	-	HETEROCYCLUS
3,4 3,8	8	MULTIPLETT	-	HETEROCYCLUS
2,5	3	SINGULETT	-	METHYLGRUPPE
2,2	3	SINGULETT	-	METHYLGRUPPE

: C-MORPHOLINO-N-(4-ETHOXY-1-NAPHTHYL)-FORMIMIDOYLCHLORID
SCHMELZPUNKT: 103 GRAD C
N M R - S P E K T R U M (60 MHz):
STANDARD: TETRAMETHYLSILAN
TEMPERATUR: 30 GRAD C
LOESUNGSM.: DEUTEROCHLOROFORM

ppm	H	STRUKTUR	J (Hz)	ZUORDNUNG
6,4 8,5	6	MULTIPLETT	-	AROMAT
3,9 4,2	2	QUARTETT	-	METHYLENGRUPPE
3,8	8	SINGULETT	-	ALIPHAT
1,2 1,6	3	TRIPLETT	-	METHYLGRUPPE

6.2.3. Reaktionsgleichungen

Ist die Nummer bekannt, unter der eine Reaktion im System abgespeichert ist, so kann man mit

G,R_a

alle gespeicherten Daten zu dieser Reaktionsgleichung abfragen.

Beispiel:

```
g,r 1685
 1685:  1607(1) + 2286(2) + 2375(2)  =  1608(1) + 2296(1) + 2299(1)
      1607:  TRIAETHYLAMIN
      2286:  2.2.2-TRICHLOR-N-CYANACETAMIDIN
      2375:  N-(p-PHENYL-PHENYL)-MORPHOLIN-4-CARBIMIDOYLCHLORID
      1608:  TRIAETHYLAMINHYDROCHLORID
      2296:  2.2.2-TRICHLOR-N-(2-MORPHOLINO-6-PHENYL-4-CHINAZOLINYL)ACETAM
             -HYDROCHLORID
      2299:  2-IMINO-6-MORPHOLINO-1-(p-PHENYLEN)-(TRICHLORMETHYL)-1.2-DIHY
             1.3.5-TRIAZIN
    AUSB.:   2296  =  18,0 %
             2299  =   9,0 %
    TEMP.:   80 GRAD C
    R-ART:   98
```

In dieser lexikographischen Ausführlichkeit benötigt man die Daten allerdings nur selten, wichtiger sind die Abfragen nach Teilmengen der gespeicherten Daten:

RB,R_a

Hiermit werden die Reaktionsbedingungen einer Reaktion auf den Bildschirm eingespielt.

Beispiel:

```
rb,r 1667
 1667:  1607(1) + 1972(1) + 2379(1)  =  1608(1) + 2278(1)
     AUSB.:   2278  =  27,0 %
     TEMP.:   82 GRAD C
     DRUCK:   1 ATM
     LOES.:   DIMETHOXYETHAN
```

Mit dem obigen Code werden nur die Reaktionsbedingungen, die unter der Reaktionsnummer gespeichert sind, abgerufen. Die Verbindungen werden durch die Verbindungsnummer angezeigt.
Mit der Erweiterung des Codes

RB,V,R_a

werden die Namen der Verbindungen auf den Bildschirm geschrieben:

Beispiel:

```
rb,v,r 1667
 1667:  1607(1) + 1972(1) + 2379(1)  =  1608(1) + 2278(1)
      1607:  TRIAETHYLAMIN
      1972:  ANILIN
      2379:  4-CHLOR-2-MORPHOLINCHINAZOLIN
      1608:  TRIAETHYLAMINHYDROCHLORID
      2278:  2-MORPHOLINO-4-(N-PHENYL-AMINO)CHINAZOLIN
     AUSB.:   2278  =  27,0 %
     TEMP.:   82 GRAD C
     DRUCK:   1 ATM
     LOES.:   DIMETHOXYETHAN
```

Die umgekehrte Abfrage, nach allen Reaktionen, die bei einer bestimmten Temperatur ablaufen, ist ebenfalls codiert:

R,T_a

Beispiel:

```
r,t 20

   87:  6(1) + 1134(2)  =  1149(2)
      TEMP.:   20 GRAD C

 1388:  1111(1) + 1783(1) + 1786(1)  =  1108(1) + 1784(1) + 1785(1)
      TEMP.:   20 GRAD C

 1393:  1111(1) + 1783(1) + 1794(1)  =  1108(1) + 1785(1) + 1793(1)
      TEMP.:   20 GRAD C

 1395:  5(1) + 2218(1)  =  2217(1)
      TEMP.:   20 GRAD C

 1396:  1111(1) + 1783(1) + 1800(1)  =  1108(1) + 1785(1) + 1804(1)
      TEMP.:   20 GRAD C

 1397:  1111(1) + 1783(1) + 1802(1)  =  1108(1) + 1785(1) + 1806(1)
      TEMP.:   20 GRAD C

 1400:  1110(1) + 1612(1)  =  7(1) + 1788(1)
      TEMP.:   0 - 20 GRAD C

 1401:  2(2) + 1808(1)  =  1809(1)
      TEMP.:   20 GRAD C

 1402:  2(2) + 1810(1)  =  1811(1)
      TEMP.:   20 GRAD C

 1404:  1110(1) + 1814(1)  =  7(1) + 1815(1)
      TEMP.:   0 - 20 GRAD C
AUSGABE FORTSETZEN?  (J;N)
j

 1406:  1110(1) + 1816(1)  =  1817(1) + 1818(1)
      TEMP.:   20 GRAD C

 1408:  51(2) + 1821(1)  =  1823(1)
      TEMP.:   20 GRAD C

 1409:  2(2) + 1824(1)  =  1825(1)
      TEMP.:   20 GRAD C
```

Ein (beliebig großes) Temperaturintervall kann vom Benutzer vorgegeben werden:

R,T_a_1; a_2

Beispiel:

```
r,t 30;50

   21:  1(1) + 1108(1)  =  2(1) + 1111(1)
      TEMP.:   30 GRAD C

  286:  8(1) + 52(1) + 1269(1)  =  3(1) + 61(1) + 1268(1)
      TEMP.:   50 GRAD C

  585:  36(1) + 1204(1)  =  56(1) + 1406(1)
      TEMP.:   40 GRAD C

  609:  46(1) + 1390(1)  =  7(1) + 19(1) + 1412(1)
      TEMP.:   40 GRAD C

  648:  51(2) + 1263(1)  =  1415(1)
      TEMP.:   33 GRAD C

  656:  3(1) + 1417(1)  =  1416(1)
      TEMP.:   30 GRAD C

  701:  3(1) + 7(1) + 1397(1)  =  54(1) + 1181(1)
      TEMP.:   50 GRAD C

  744:  5(3) + 1423(1)  =  3(2) + 1425(1)
      TEMP.:   50 GRAD C

 1441:  81(2) + 1135(1) + 1288(1) + 1886(1)  =  93(1) + 1135(1) + 1887(1)
      TEMP.:   50 GRAD C

 1443:  1(1) + 1888(1)  =  1889(1)
      TEMP.:   20 - 40 GRAD C
AUSGABE FORTSETZEN?  (J;N)
j

 1464:  32(1) + 1779(1)  =  3(2) + 2240(1)
      TEMP.:   30 - 140 GRAD C

 1472:  32(1) + 2096(1)  =  3(2) + 2241(1)
      TEMP.:   30 - 140 GRAD C

 1476:  32(1) + 2221(1)  =  3(2) + 2242(1)
      TEMP.:   30 - 140 GRAD C

 1486:  1110(1) + 1327(1) + 1607(2) + 1979(1)  =  7(1) + 1608(2) + 1980(1
      TEMP.:   50 GRAD C

 1487:  1110(1) + 1325(1) + 1607(2) + 1979(1)  =  7(1) + 1608(2) + 1981(1
      TEMP.:   50 GRAD C
```

Die Anzeige der Verbindungsnamen bei Temperaturabfragen erfolgt durch diese Codes:

V,R,T_a

bzw. bei Temperaturintervallen:

V,R,T_a_1; a_2

Beispiele:

,r,t 20

```
  87:  6(1) + 1134(2)  =  1149(2)
       6:  SAUERSTOFF
    1134:  AETHYLEN
    1149:  ACETALDEHYD
   TEMP.:   20 GRAD C

1388:  1111(1) + 1783(1) + 1786(1)  =  1108(1) + 1784(1) + 1785(1)
    1111:  TETRACHLORKOHLENSTOFF
    1783:  TRIPHENYLPHOSPHAN
    1786:  MORPHOLIN-4-CARBONSAEURE-trans-STYRYLAMID
    1108:  CHLOROFORM
    1784:  N-trans-STYRYL-MORPHOLIN-4-CARBIMIDOYLCHLORID
    1785:  TRIPHENYLPHOSPHANOXID
   TEMP.:   20 GRAD C

1393:  1111(1) + 1783(1) + 1794(1)  =  1108(1) + 1785(1) + 1793(1)
    1111:  TETRACHLORKOHLENSTOFF
    1783:  TRIPHENYLPHOSPHAN
    1794:  MORPHOLIN-4-CARBONSAEURE-1.2-DIPHENYL-VINYLAMID
    1108:  CHLOROFORM
    1785:  TRIPHENYLPHOSPHANOXID
    1793:  1.2-DIPHENYL-VINYL-MORPHOLIN-4-CARBIMIDOYLCHLORID
   TEMP.:   20 GRAD C

1395:  5(1) + 2218(1)  =  2217(1)
       5:  WASSERSTOFF
    2218:  STYROL
    2217:  ETHYLBENZOL
   TEMP.:   20 GRAD C

1396:  1111(1) + 1783(1) + 1800(1)  =  1108(1) + 1785(1) + 1804(1)
    1111:  TETRACHLORKOHLENSTOFF
    1783:  TRIPHENYLPHOSPHAN
    1800:  (3.4.5-TRIMETHOXYPHENYL)-CARBAMIDSAEUREMORPHOLID
    1108:  CHLOROFORM
    1785:  TRIPHENYLPHOSPHANOXID
    1804:  C-MORPHOLINO-N-(3.4.5-TRIMETHOXYPHENYL)-FORMIMIDOYLCHLORID
   TEMP.:   20 GRAD C
```

...

Fortsetzung des Beispiels "V,R,T 2o"

```
 1397:  1111(1) + 1783(1) + 1802(1)  =  1108(1) + 1785(1) + 1806(1)
      1111:  TETRACHLORKOHLENSTOFF
      1783:  TRIPHENYLPHOSPHAN
      1802:  (4-ETHOXY-1-NAPHTHYL)-CARBAMIDSAEUREMORPHOLID
      1108:  CHLOROFORM
      1785:  TRIPHENYLPHOSPHANOXID
      1806:  C-MORPHOLINO-N-(4-ETHOXY-1-NAPHTHYL)-FORMIMIDOYLCHLORID
     TEMP.:   20 GRAD C

 1400:  1110(1) + 1612(1)  =  7(1) + 1788(1)
      1110:  PHOSGEN
      1612:  N.N'-DICYCLOHEXYLHARNSTOFF
         7:  KOHLENDIOXID
      1788:  N.N'-DICYCLOHEXYL-CHLORFORMAMIDIN-HYDROCHLORID
     TEMP.:   0 - 20 GRAD C

 1401:  2(2) + 1808(1)  =  1809(1)
         2:  CHLORWASSERSTOFF(SAEURE)
      1808:  DIPHENYLCARBODIIMID
      1809:  N.N'-DIPHENYL-CHLORFORMAMIDIN-HYDROCHLORID
     TEMP.:   20 GRAD C

 1402:  2 2) + 1810(1)  =  1811(1)
         2:  CHLORWASSERSTOFF(SAEURE)
      1810:  DI-n-BUTYLCARBODIIMID
      1811:  N.N'-DI-n-BUTYLCHLORFORMAMIDINHYDROCHLORID
     TEMP.:   20 GRAD C

 1404:  1110(1) + 1814(1)  =  7(1) + 1815(1)
      1110:  PHOSGEN
      1814:  N.N'-DIMETHYLHARNSTOFF
         7:  KOHLENDIOXID
      1815:  N.N'-DIMETHYL-CHLORFORMAMIDIN-HYDROCHLORID
     TEMP.:   0 - 20 GRAD C
AUSGABE FORTSETZEN?  (J;N)
n
```

Der Druck, unter dem eine Reaktion abläuft, wird mit

P,R_a

erfragt, die Erweiterung mit Verbindungsnamen lautet:

P,V,R_a

Beispiel:

```
p,r 1624
 1624:   3(l) + 1977(l)   =   2(l) + 2168(l)
      DRUCK:    1 ATM
```

Das Lösungsmittel, das bei einer Umsetzung verwendet wird, erfragt man mit

L,R_a

bzw.

L,V,R_a

Dabei ist zu beachten, daß auch Lösungsmittelgemische angezeigt werden können.

Beispiele:

l,r 1625

```
1625:   3 2) + 1977(1)   =   2(1) + 7(1) + 1368 1) + 2167 1)
LOES.:    WASSER
          METHYLALKOHOL
```

l,v,r 1624

```
1624:   3(1) + 1977(1)   =   2(1) + 2168(1)
   3:   WASSER
1977:   3-CHLOR-N- CHLOR DIMETHYLAMINO)-METHYLEN)-6-METHYL-BENZO(b)TH
        EN-2-CARBOXAMID
   2:   CHLORWASSERSTOFF(SAEURE)
2168:   3-CHLOR-N-DIMETHYLCARBAMOYL-6-METHYLBENZO(b)THIOPHEN-2-CARBOX
LOES.:    WASSER
          DIMETHYLFORMAMID
```

Nach Katalysatoren kann man auf zwei Arten fragen.
Einmal nach dem schon bekannten Abfrageschema:

KAT,R_a

bzw.

KAT,V,R_a

oder auch, wenn ein Stoff oder eine Stoffkombination
auf die Verwendung als Katalysator überprüft werden soll:

R,KAT_a

oder V,R,KAT_a

Beispiele:

```
t,r 1628
628:  32(2) + 2181(1)  =  3(4) + 2184(1)
    KAT.:   TRIETHYLAMIN/PYRIDIN
```

```
t,v,r 1627
627:  32(2) + 2180(1)  =  3(4) + 2183(1)
      32:  THIONYLCHLORID
    2180:  PHENYLEN-1.2-DIACRYLSAEURE
       3:  WASSER
    2183:  3.4-DICHLORBENZO(1.2-b:4.3-b')DITHIOPHEN-2.5-DICARBONYLCHLORID
    KAT.:   TRIETHYLAMIN/PYRIDIN
```

Die Ausbeute einer Reaktion stellt ein wichtiges Beschreibungsmerkmal dar, man erfragt sie mit

A,R_a

Beispiel:

```
a,r 1645
  1645:  97(3) + 1774(3)  =  98(3) + 2211(1) + 2213(1)
       AUSB.:    2211   =  30,0 -  40,0 %
                 2213   =  20,0 -  30,0 %
```

Die erweiterte Abfragemöglichkeit nach Ausbeuten und Verbindungsnamen ist ebenfalls programmiert:

Beispiel:

```
a,v,r 1645
  1645:  97(3) + 1774(3)  =  98(3) + 2211(1) + 2213(1)
         97:  SILBERCYANAT
       1774:  N-p-CHLOR-PHENYL-MORPHOLIN-4-CARBIMIDOYLCHLORID
         98:  SILBERCHLORID
       2211:  2-MORPHOLINO-6-CHLOR-CHINAZOLON-(4)
       2213:  1-(4-CHLORPHENYL)-2-MORPHOLINO-5-(4-CHLORPHENYLIMINO-MORPHOLIN
              ETHYL)-1.3.5-TRIAZINDION-(4.6)
      AUSB.:   2211   =  30,0 -  40,0 %
               2213   =  20,0 -  30,0 %
```

Nach der Temperatur einer Reaktion fragt man mit

T,R_a

oder mit

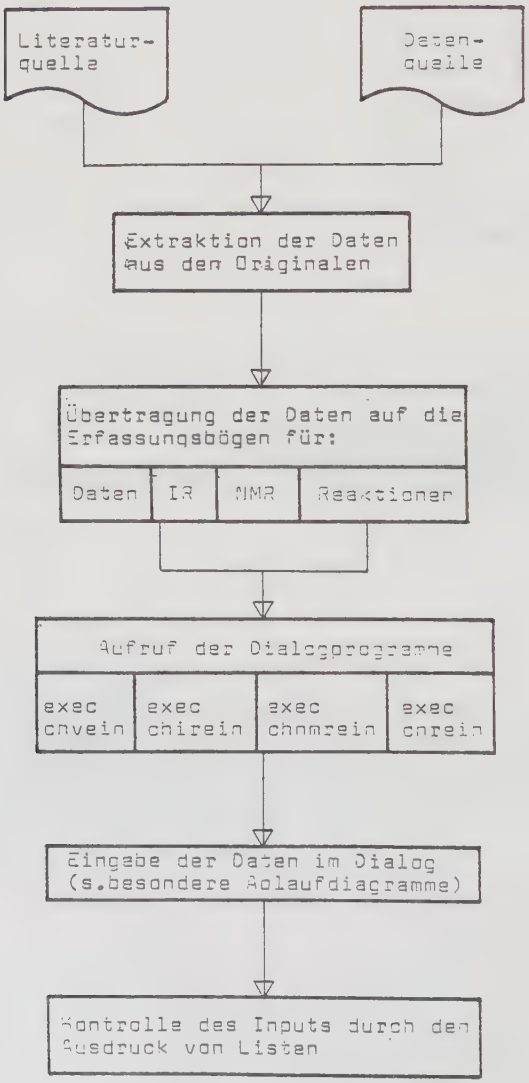
T,V,R_a

Beispiel:

```
t,r 1648
 1648:  1114(1) + 1607(2) + 1772(2)  =  1608(2) + 2261(1)
     TEMP.:   82 GRAD C


t,v,r 1636
 1636:  1607(1) + 1769(1) + 2194(1)  =  1608(1) + 2195(1)
      1607:  TRIAETHYLAMIN
      1769:  N-PHENYL-MORPHOLIN-4-CARBIMIDOYLCHLORID
      2194:  CARBHYDRAZIDSAEUREETHYLESTER
      1608:  TRIAETHYLAMINHYDROCHLORID
      2195:  PHENYLIMINO-MORPHOLINO-(2-CARBETHOXYHYDRAZINO)-METHAN
     TEMP.:   20 GRAD C
```

6.3. Eingabe

Darstellung des verallgemeinerten Inputs im Ablaufdiagramm



Die auf den Seiten 234, 239, 245 und 248 abgebildeten Druckerprotokolle geben den im Ablaufdiagramm skizzierten Eingabevorgang wieder.
Sie stellen den Versuch dar, mit den zur Verfügung stenenden Mittel eine möglichst benutzerfreundliche Dateneingabe zu programmieren. Da es bei dieser Arbeit auf die grundsätzliche Lösung des Problems und nicht auf die Verwendung der perfektesten Datenverarbeitungstechnik ankommt, wurde in der Experimentalphase der Systementwicklung auf die aufwendige Realisierung der Menütechnik verzichtet. (Das auszufüllende Eingabeformular erscheint auf dem Bildschirm, der Benutzer wird durch den Cursor zu den auszufüllenden Kategorien geführt.)

6.3.1. Eigenschaften von Verbindungen

Am Beispiel verschiedener Publikationen wird im folgenden Kapitel gezeigt, wie der Literaturauswerte- und Einspeichervorgang aus der Sicht des Systembetreibers abläuft.
Aus der Veröffentlichung (105) werden die Daten eines bifunktionellen Chlorformamidins entnommen und in den Erfassungsbogen übertragen.
Die Verbindungsnummer wird fortlaufend vergeben.
Das System weist schon mit Daten belegte Nummern zurück.
Der Name wird aus der Literatur übernommen.
Die Reihenfolge der Elemente bei der Eingabe der Summenformel ist nicht festgelegt, durch das Programm werden die Elemente nach dem Hill'schen System intern sortiert und erscheinen bei der Ausgabe der Daten in der richtigen Reihenfolge.
Schmelzpunkte werden mit der Einheit "°C" direkt aus der Literatur excerpiert, die Angabe einer Dezimalstelle ist möglich.
An Bemerkungen zum Schmelzpunkt stehen zur Verfügung:

z - für Zersetzung und
ö - für Öl

Die Siedepunkte und die Kommentare dazu entsprechen in der formalen Behandlung dem Verfahren bei den Schmelzpunkten.
Der Druck beim Siedepunkt eines Stoffes kann in Torr oder in Atmosphären angegeben werden.

Die Kennziffer für die Familienzugehörigkeit einer Verbindung bietet folgende Möglichkeiten:

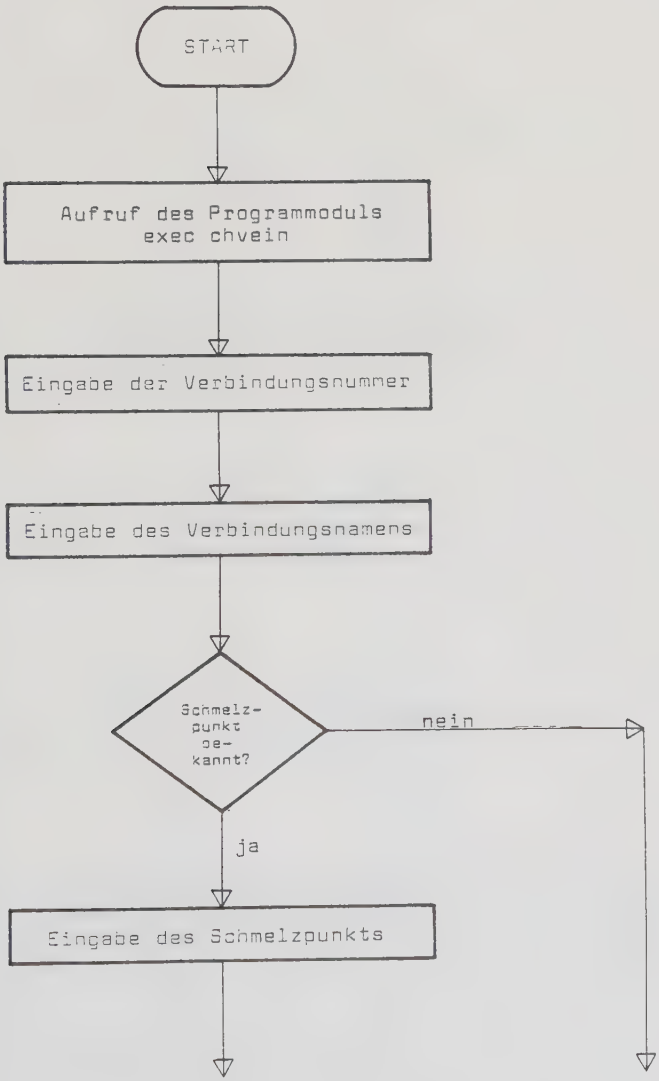
F 1oo - Peptide
F 2oo - Säuren
F 3oo - Halogenide
F 4oo - Carbonsäureester
F 5oo - Olefine
F 6oo - Harnstoffe
F 61o - Chlorformamidine
F 62o - Isocyaniddichloride

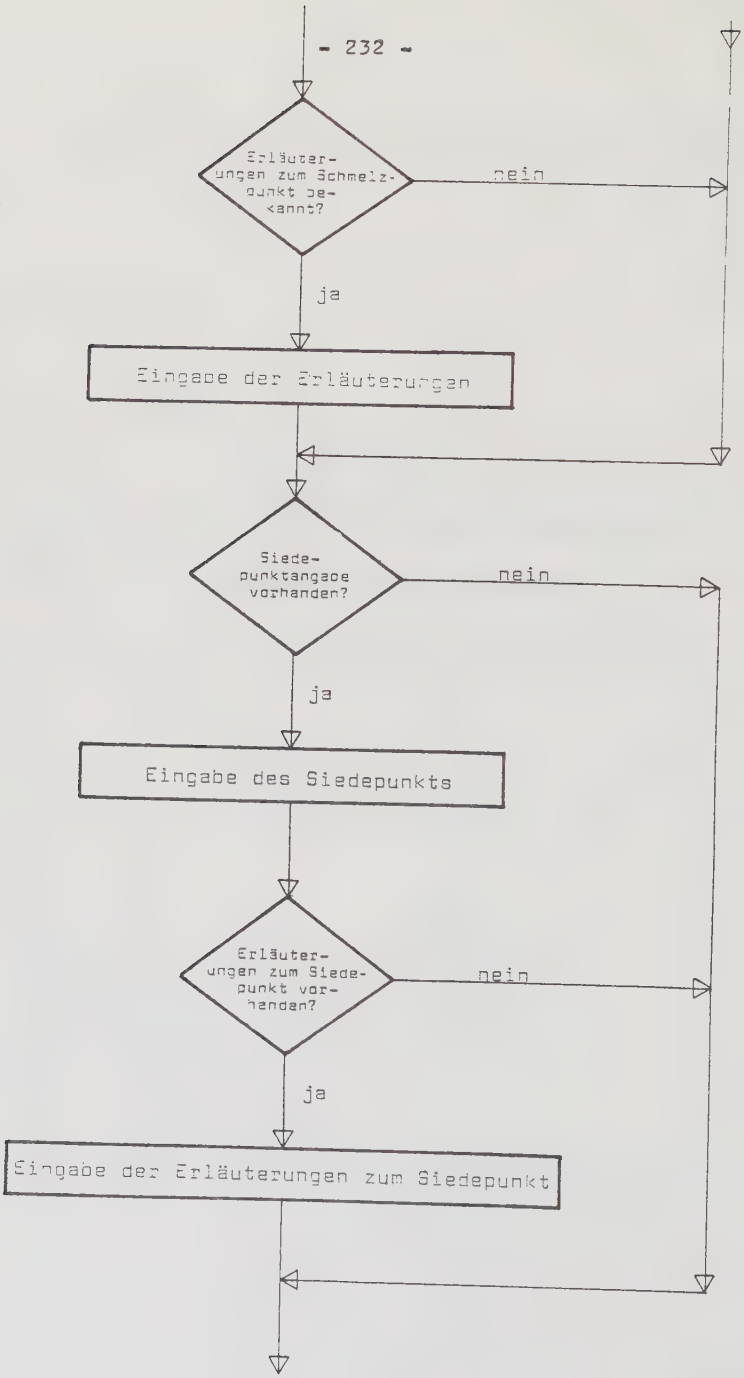
Die Basenkonstante kann auf zwei Stellen hinter dem Komma genau eingespeichert werden.

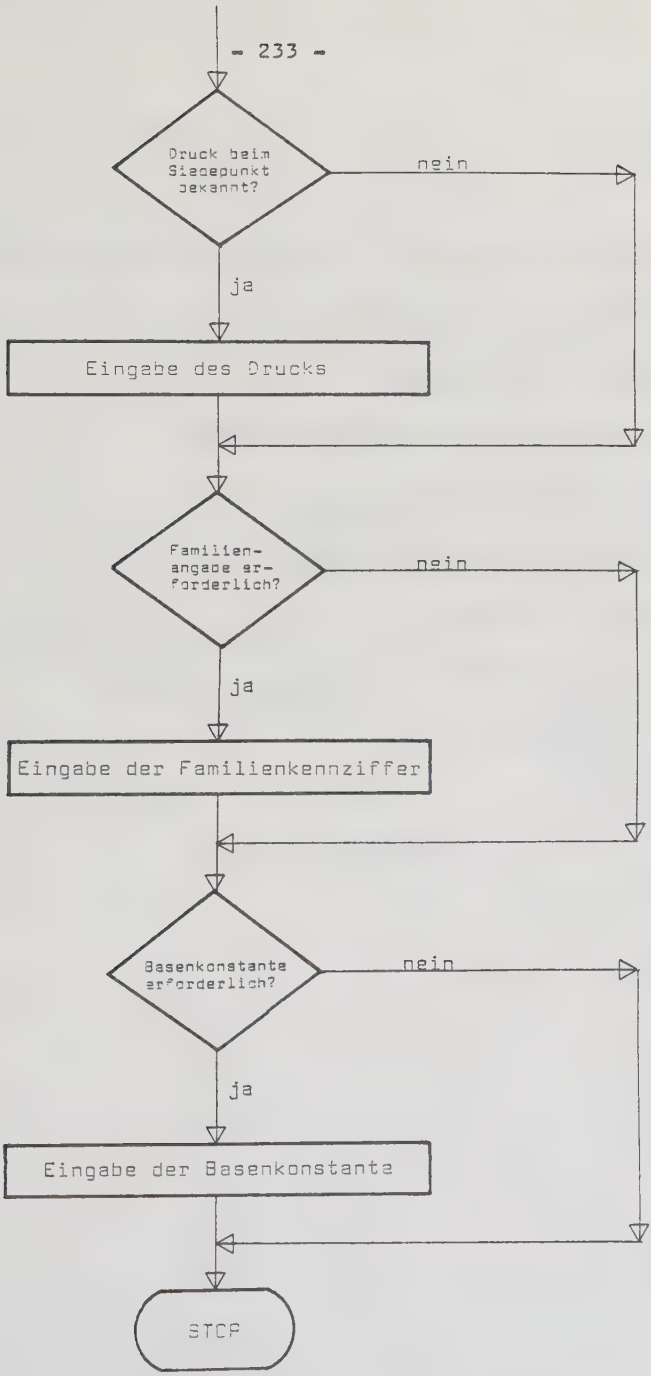
Beispiel aus (105):

exec chvein				
Nr.:	2078			
Name:	N,N'-bis (chlormorpholino-methylen)-p-phenylendiamin			
Summenformel:	$C_{16}H_{20}Cl_2N_4O_2$			
Fp.:	232	°C		
Bem. Fp.:	—			
Kp.:	—	°C		
Bem. Kp.:	—			
Druck Kp.:	—	atm		Torr
Familie:	610			
Basenkonstante:	—			

Ablaufdiagramm für die Eingabe von Daten zu Verbindungen




Erläuterungen zum Schmelzpunkt bekannt?
nein
ja
Eingabe der Erläuterungen
Siedepunktangabe vorhanden?
nein
ja
Eingabe des Siedepunkts
Erläuterungen zum Siedepunkt vorhanden?
nein
ja
Eingabe der Erläuterungen zum Siedepunkt


Druck beim Siedepunkt bekannt?
nein
ja
Eingabe des Drucks
Familien-angabe er-forderlich?
nein
ja
Eingabe der Familienkennziffer
Basenkonstante erforderlich?
nein
ja
Eingabe der Basenkonstante
STOP

Druckerprotokoll der Eingabe von Verbindungsdaten:

```
exec chvein
ENTHALTEN DIE EINZUGEBENDEN VERBINDUNGEN FAMILIENANGABEN?   (J;N
j
BITTE NUMMER UND NAME EINGEBEN
2078 n.n'->bis<-(chlor-morpholino-methylen)->p<-phenylendiamin
BITTE SUMMENFORMEL EINGEBEN
c16h20n4o2cl2
BITTE SCHMELZPUNKT EINGEBEN
232
BITTE BEMERKUNGEN ZUM SCHMELZPUNKT EINGEBEN

BITTE SIEDEPUNKT EINGEBEN

BITTE BASIZITAETSKONSTANTE EINGEBEN

BITTE ACIDITAETSKONSTANTE EINGEBEN

BITTE FAMILIE EINGEBEN
610
 2078:  N.N'-bis-(CHLOR-MORPHOLINO-METHYLEN)-p-PHENYLENDIAMIN
        C 16, H 20, N 4, O 2, Cl 2
        FAMILIE:              610 CHLORFORMAMIDIN
        SCHMELZPUNKT:         232 GRAD C
        SIEDEPUNKT:           --
        pKb:                  --
        pKs:                  --
KORREKT?   (J;N)
j
BITTE NUMMER UND NAME EINGEBEN
ende
LISTE GEWUENSCHT?   (J;N)
n
READY
```

5.3.2. Spektraldaten

Das Infrarotspektrum der Verbindung mit der CHEMIS-Nummer 2078 wird unter dieser Nummer gespeichert, bei zeitlichem Auseinanderliegen von Spektren- und Verbindungseingabe werden die Daten von dem Rechner zusammengeführt.
Die Bezeichnung des Meßgeräts erfolgt durch die Codes auf S.91.
Das Medium wird durch Eingabe der Verbindungsnummer verschlüsselt.
Die Kennzahlen für die Schwingungselemente und Interpretationen werden aus den Listen auf S.95 und S.94 entnommen.
Die Kennzeichnung der Bandenintensität wird durch den Intensitätsschlüssel vorgenommen.

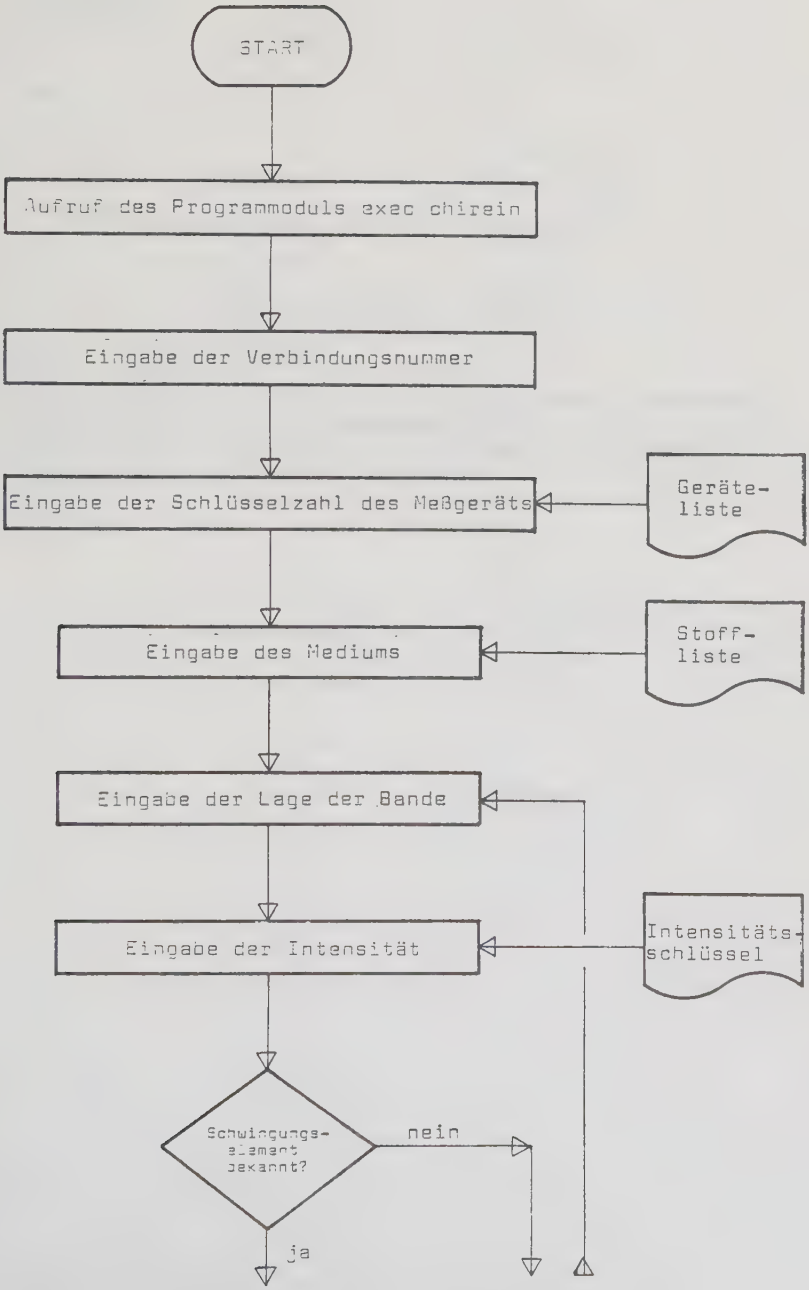
Intensitätsschlüssel

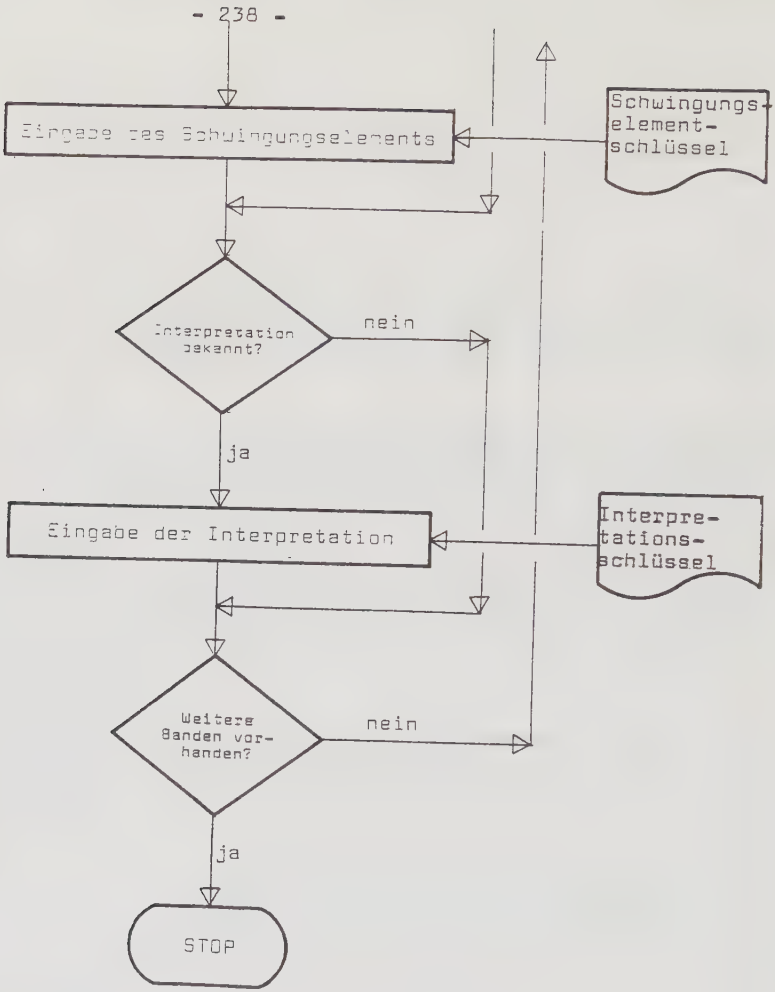
1 = sehr stark
2 = stark
3 = mittel
4 = schwach

Beispiel aus (105):

exec chirein				
Nr.: 2078		Gerät: 2		Medium: 53
	Lage	Intensität	Element	Interpretation
1	1645	1	60	473
2				
3				
4				
5				
6				
7				
8				
9				
10				

Ablaufdiagramm für die Eingabe von Infrarotspektren




Eingabe des Schwingungselements
Schwingungs-element-schlüssel
Interpretation bekannt?
nein
ja
Eingabe der Interpretation
Interpre-tations-schlüssel
Weitere Banden vor-handen?
nein
ja
STOP

Druckerprotokoll der Eingabe des IR-Spektrums:

```
ec stke.clist(chirein)

TTE VERBINDUNGSNUMMER, NR. DES MEDIUMS UND NR. DES MESSGERAETES EINGEBEN
78 53 1
TTE BANDE 1 EINGEBEN (WELLENZAHL, INTENS., SCHW.-ELEMENT, INTERPRETATION)
45 1 60 473
TTE BANDE 2 EINGEBEN (WELLENZAHL, INTENS., SCHW.-ELEMENT, INTERPRETATION)

TES IR-SPEKTRUM UEBERSCHREIBEN?   (J,N)

078:  N.N'-bis-(CHLOR-MORPHOLINO-METHYLEN)-p-PHENYLENDIAMIN
      I R - S P E K T R U M   (PERKIN-ELMER 177):
      MEDIUM:        KALIUMBROMID
      WELLENZAHL   INTENSITAET   ZUORDNUNG
        1645        SEHR STARK   C-N-DOPPELBINDUNGSVALENZSCHWINGUNG
                                 CHLORFORMAMIDIN

TTE VERBINDUNGSNUMMER, NR. DES MEDIUMS UND NR. DES MESSGERAETES EINGEBEN
ie
STE GEWUENSCHT?   (J;N)

ADY
```

Das NMR-Spektrum eines Stoffes wird durchnumeriert, das erste einzugebende Spektrum sollte das mit den Standardmeßbedingungen sein, das zweite dokumentiert die Änderungen der Temperatur oder des Lösungsmittels.
Die Meßfrequenz wird in der Einheit "MHz" nachgewiesen.
Standard und Lösungsmittel erhalten die Schlüsselzahlen für chemische Verbindungen.
Die Temperatur hat die Maßeinheit "°C".
Bei der Angabe der Peaklage in ppm ist eine Stelle hinter dem Komma erlaubt (Auflösung 0,1 ppm), Bereiche können verwendet werden.
Die Signalintensität wird duch die Anzahl der H-Atome charakterisiert.
Für die Aufspaltung verwendet man folgende Abkürzungen:

s = Singulett
d = Dublett
t = Triplett
m = Multiplett

Die Eingabe der Kopplungskonstante wird bei Singuletts unterdrückt, sonst ist die Angabe der Größe in "Hz" möglich.
Die Zuordnung erfolgt durch die Verwendung der Schlüsselzahlen für Interpretationen. (s.S.95)

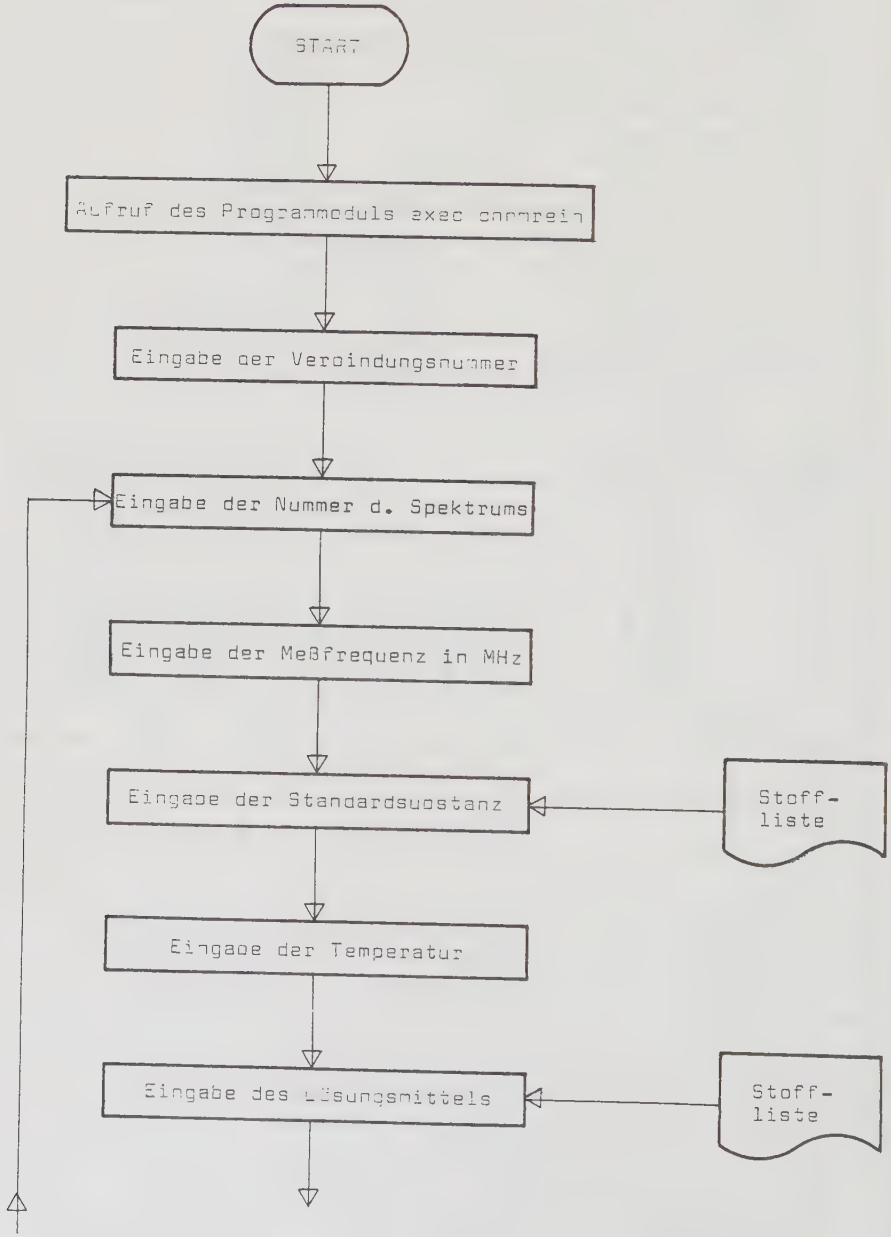
Beispiel aus (87):

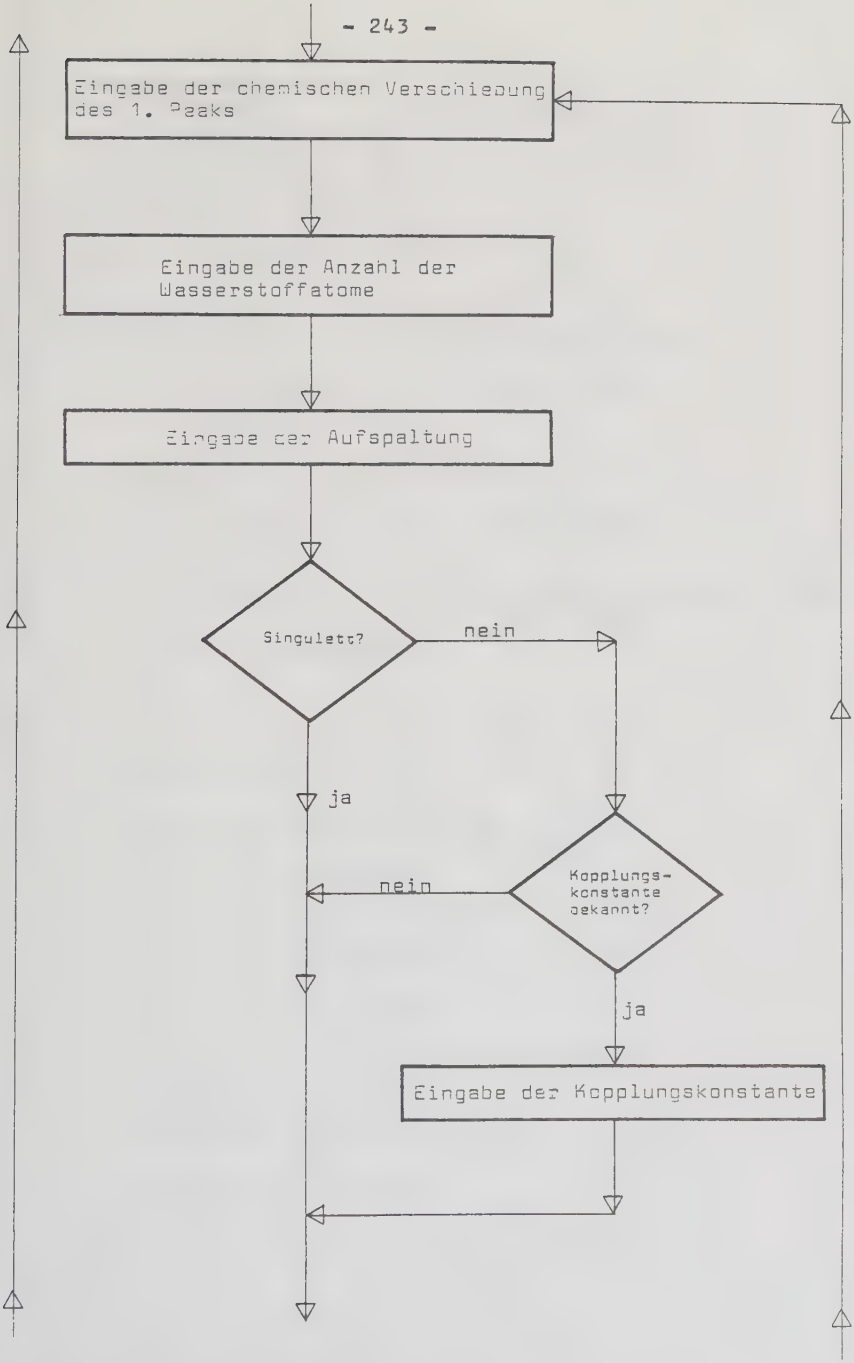
exec chnmrein			
Verbindungsnummer:	1784		
Nr. des Spektrums:	1		
Messfrequenz:	60	MHz	
Standard:	1960		
Temperatur:	30	°C	
Lösungsmittel:	1926		

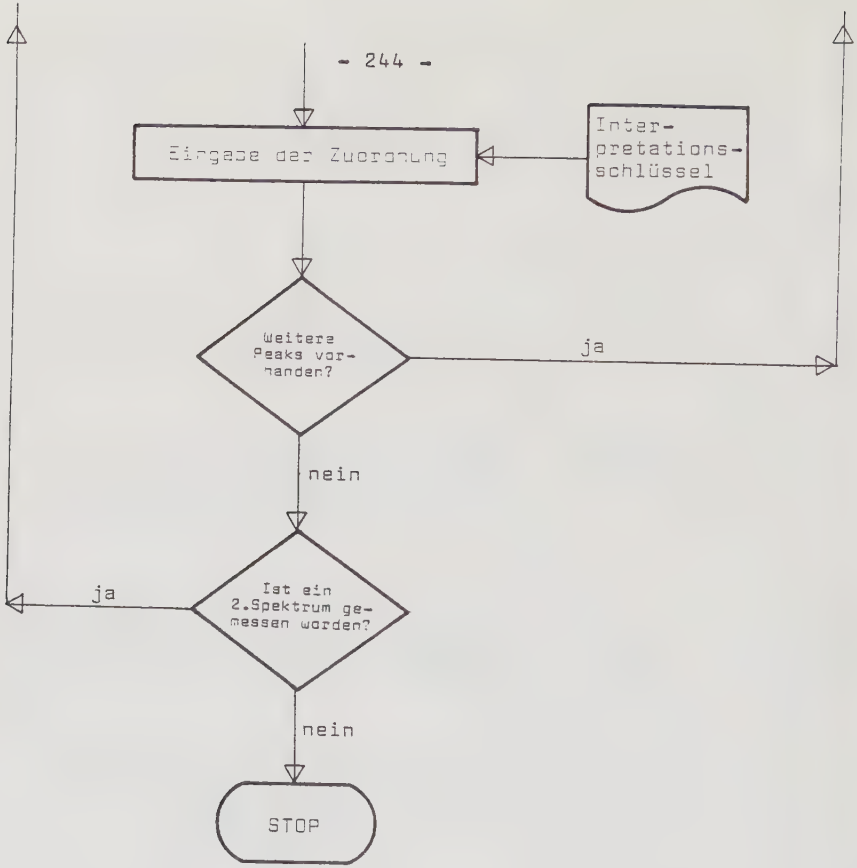
Peak	δ [ppm]	H	Aufspaltung	Hz	Zuordnung
1	7,5	1	d	—	130
2	7,3	5	m	—	110
3	6,5	1	d	—	130
4	3,6	8	s	—	100
5					
6					
7					
8					
9					
10					

je

Ablaufdiagramm für die Eingabe von NMR-Spektren



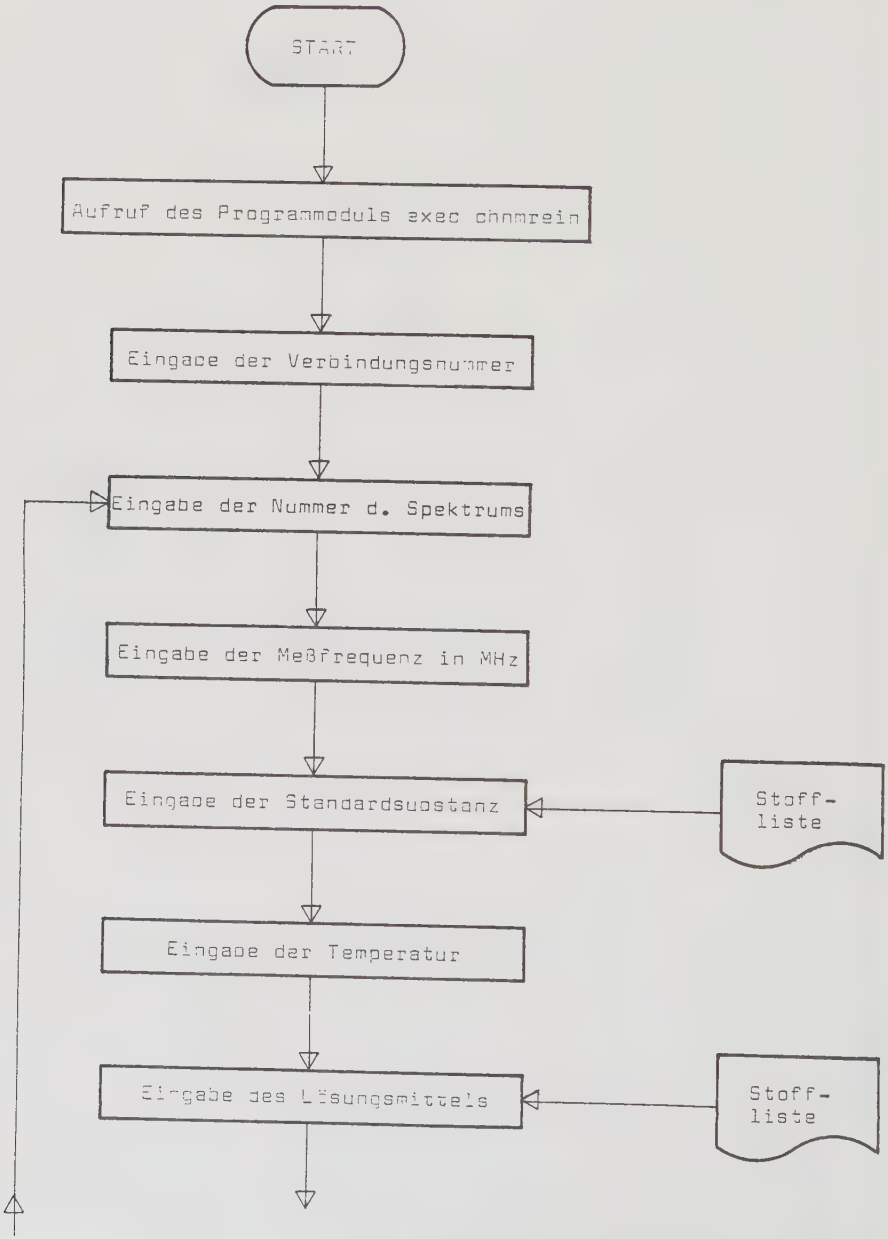

Eingabe der chemischen Verschiebung des 1. Peaks
Eingabe der Anzahl der Wasserstoffatome
Eingabe der Aufspaltung
Singulett?
nein
ja
Kopplungs-konstante bekannt?
nein
ja
Eingabe der Kopplungskonstante

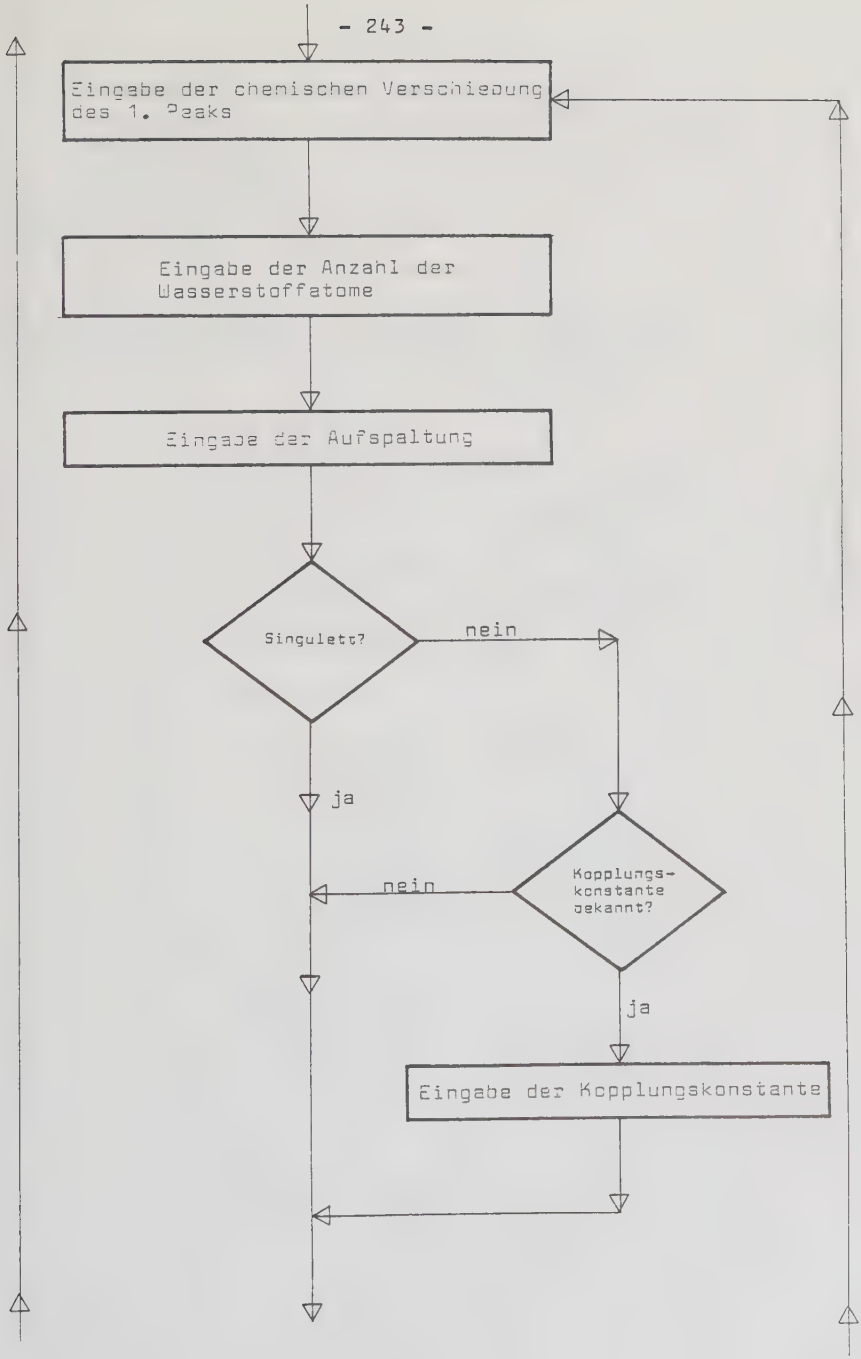

Eingabe der Zuordnung
Inter-
pretations-
schlüssel
Weitere
Peaks vor-
handen?
ja
nein
Ist ein
2.Spektrum ge-
messen worden?
ja
nein
STOP

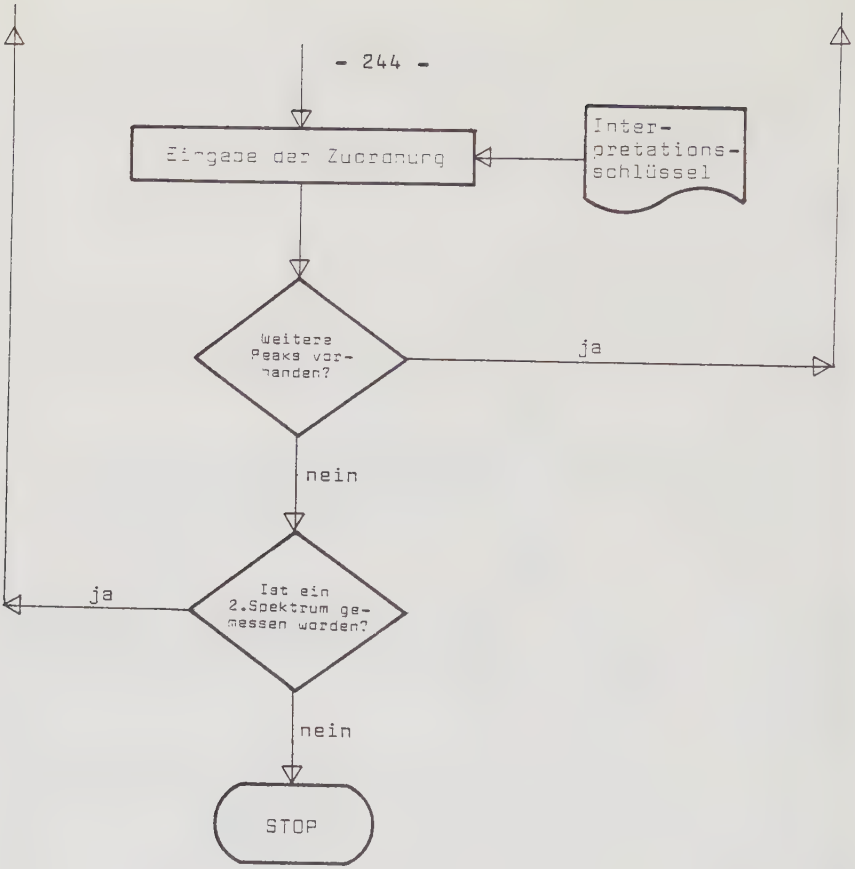
Druckerprotokoll der Eingabe eines NMR-Spektrums:

```
exec stke.clist(chnmrein)

BITTE VERBINDUNGSNUMMER EINGEBEN
1784
NMR-SPEKTRUM UEBERSCHREIBEN BZW. LOESCHEN?   (J;N)
j
1. MESSBEDINGUNGEN:
BITTE FREQUENZ, STANDARD, TEMPERATUR U. LOESUNGSMITTEL EINGEBEN
60 1960 30 1926
PEAK 1:  BITTE CHEM. VERSCHIEBUNG (ppm) EINGEBEN
7,5
PEAK 1:  BITTE ANZAHL H EINGEBEN
1
PEAK 1:  BITTE SIGNAL-STRUKTUR EINGEBEN   (S;D;T;Q;M)
d
PEAK 1:  BITTE KOPPLUNGSKONSTANTE J (Hz) EINGEBEN

PEAK 1:  BITTE ZUORDNUNG EINGEBEN
130
PEAK 2:  BITTE CHEM. VERSCHIEBUNG (ppm) EINGEBEN
7,3
PEAK 2:  BITTE ANZAHL H EINGEBEN
5
PEAK 2:  BITTE SIGNAL-STRUKTUR EINGEBEN   (S;D;T;Q;M)
m
PEAK 2:  BITTE KOPPLUNGSKONSTANTE J (Hz) EINGEBEN

PEAK 2:  BITTE ZUORDNUNG EINGEBEN
110
PEAK 3:  BITTE CHEM. VERSCHIEBUNG (ppm) EINGEBEN
6,5
PEAK 3:  BITTE ANZAHL H EINGEBEN
1
PEAK 3:  BITTE SIGNAL-STRUKTUR EINGEBEN   (S;D;T;Q;M)
d
PEAK 3:  BITTE KOPPLUNGSKONSTANTE J (Hz) EINGEBEN

PEAK 3:  BITTE ZUORDNUNG EINGEBEN
130
PEAK 4:  BITTE CHEM. VERSCHIEBUNG (ppm) EINGEBEN
3,6
PEAK 4:  BITTE ANZAHL H EINGEBEN
8
PEAK 4:  BITTE SIGNAL-STRUKTUR EINGEBEN   (S;D;T;Q;M)
s
PEAK 4:  BITTE ZUORDNUNG EINGEBEN
100
PEAK 5:  BITTE CHEM. VERSCHIEBUNG (ppm) EINGEBEN
```

Ablaufdiagramm für die Eingabe von NMR-Spektren




Eingabe der chemischen Verschiebung des 1. Peaks
Eingabe der Anzahl der Wasserstoffatome
Eingabe der Aufspaltung
Singulett?
nein
ja
Kopplungs-konstante bekannt?
nein
ja
Eingabe der Kopplungskonstante


Eingabe der Zuordnung
Inter-
pretations-
schlüssel
Weitere
Peaks vor-
handen?
ja
nein
Ist ein
2.Spektrum ge-
messen worden?
ja
nein
STOP

Druckerprotokoll der Eingabe eines NMR-Spektrums:

```
exec stke.clist(chnmrein)

BITTE VERBINDUNGSNUMMER EINGEBEN
1784
NMR-SPEKTRUM UEBERSCHREIBEN BZW. LOESCHEN?   (J;N)
j
1. MESSBEDINGUNGEN:
BITTE FREQUENZ, STANDARD, TEMPERATUR U. LOESUNGSMITTEL EINGEBEN
60 1960 30 1926
PEAK 1:  BITTE CHEM. VERSCHIEBUNG (ppm) EINGEBEN
7,5
PEAK 1:  BITTE ANZAHL H EINGEBEN
1
PEAK 1:  BITTE SIGNAL-STRUKTUR EINGEBEN   (S;D;T;Q;M)
d
PEAK 1:  BITTE KOPPLUNGSKONSTANTE J (Hz) EINGEBEN

PEAK 1:  BITTE ZUORDNUNG EINGEBEN
130
PEAK 2:  BITTE CHEM. VERSCHIEBUNG (ppm) EINGEBEN
7,3
PEAK 2:  BITTE ANZAHL H EINGEBEN
5
PEAK 2:  BITTE SIGNAL-STRUKTUR EINGEBEN   (S;D;T;Q;M)
m
PEAK 2:  BITTE KOPPLUNGSKONSTANTE J (Hz) EINGEBEN

PEAK 2:  BITTE ZUORDNUNG EINGEBEN
110
PEAK 3:  BITTE CHEM. VERSCHIEBUNG (ppm) EINGEBEN
6,5
PEAK 3:  BITTE ANZAHL H EINGEBEN
1
PEAK 3:  BITTE SIGNAL-STRUKTUR EINGEBEN   (S;D;T;Q;M)
d
PEAK 3:  BITTE KOPPLUNGSKONSTANTE J (Hz) EINGEBEN

PEAK 3:  BITTE ZUORDNUNG EINGEBEN
130
PEAK 4:  BITTE CHEM. VERSCHIEBUNG (ppm) EINGEBEN
3,6
PEAK 4:  BITTE ANZAHL H EINGEBEN
8
PEAK 4:  BITTE SIGNAL-STRUKTUR EINGEBEN   (S;D;T;Q;M)
s
PEAK 4:  BITTE ZUORDNUNG EINGEBEN
100
PEAK 5:  BITTE CHEM. VERSCHIEBUNG (ppm) EINGEBEN
```

6.3.3. Reaktionsgleichungen

Jede chemische Reaktion ist durch eine fortlaufende Nummer gekennzeichnet.
Die Ausgangsstoffe werden als Verbindungsnummer eingegeben, in Klammern dahinter steht die Anzahl der reagierenden Moleküle.
Die Produkte werden ebenfalls durch die Verbindungsnummern beschrieben, die Anzahl der entstehenden Moleküle wird in Klammern dahinter gesetzt.
Die Ausbeute folgt unmittelbar nach dem Zahlenpaar Verbindungsnummer (Anzahl der Moleküle) und wird durch einen Schrägstrich von ihm getrennt. Es können auch Ausbeutebereiche angegeben werden.
Die Anzahl reagierender Moleküle ist auf zehn beschränkt.
Nach Eingabe des Reaktionsartenschlüssels (Liste siehe S. 112) erfolgt die Prüfung auf formale Richtigkeit der Gleichung, falsche Gleichungen werden zurückgewiesen.
Die Temperatur- oder Temperaturbereichsangaben erfolgen in "°C", beim Druck oder dem Druckbereich können sowohl Torr als auch Atmosphären verwendet werden.
Das Lösungsmittel oder das Lösungsmittelgemisch (fünf Komponenten sind erlaubt, Mischungsverhältnisse werden nicht erfaßt) wird als Verbindungsnummer verschlüsselt.
Die Schlüsselzahlen für Katalysatoren entnimmt man der Liste auf S. 114.

Beispiel aus (111):

exec chrein				
Reaktionsnummer:	1390			
Input-Partner / Anzahl der Moleküle:				
44 (1) 1790 (1)				
Output-Partner / Anzahl der Moleküle / Ausbeute:				
20 (1) 1791 (1)				
Reaktionsart:	12			
Temperatur:	0			
Druck:	1	atm	—	Torr
Lösungsmittel:	—			
Katalysator:	—			
				je

Druckerprotokoll der Eingabe einer Reaktionsgleichung:

```
exec chrein
BITTE R-NUMMER UND INPUT-PARTNER EINGEBEN
1390 44,1 1790,1
BITTE OUTPUT-PARTNER (MIT AUSBEUTE) EINGEBEN
20,1 1791,1
BITTE REAKTIONSART EINGEBEN
12
BITTE TEMPERATUR EINGEBEN
0
BITTE DRUCK EINGEBEN
1a
BITTE LOESUNGSMITTEL EINGEBEN

BITTE KATALYSATOR EINGEBEN

 1390:     44 (1) + 1790 (1)   =   20 (1) + 1791 (1)
      AUSB.:    --
      TEMP.:    0 GRAD C
      DRUCK:    1 ATM
      LOES.:    --
       KAT.:    --
      R-ART:    12
KORREKT?    (J;N)
j
BITTE R-NUMMER UND INPUT-PARTNER EINGEBEN
ende
LISTE GEWUENSCHT?    (J;N)
n
READY
```

Literaturzitate werden in CHEMIS den Reaktionen zugeordnet und auch dort eingegeben.
Für die Form der Nachweise bestehen keine besonderen bibliographischen Vorschriften, um die Zitate aussagefähiger zu machen, sollte immer der Titel der Publikation mit abgespeichert werden.
Einer Reaktion können mehrere Veröffentlichungen zugeordnet sein, sie sind dann fortlaufend durchnumeriert.

7. Zusammenfassung

Ausgehend von den Begriffen Forschung, Information, Dokumentation und Kommunikation wird mittels des Modells des Forschungsprozesses nach Popper versucht, Daten und Fakten, die der präparative organische Chemiker bei seiner Arbeit benötigt, so zu organisieren, daß möglichst wenig Verlust bei der Wiedergewinnung der Informationen eintritt.

Neben der optimalen Wiedergewinnung von Informationen ist die benutzerfreundliche und problemangepaßte Organisation der Systemsoftware ein Ziel der Arbeit. Als Untersuchungsgebiet für die Systemanalyse der Forschungsarbeit wurde die Chemie der Chlorformamidine ausgewählt.

Diese Ziele wurden durch folgende Maßnahmen erreicht:

- Die Sachverhalte (Verbindungsnamen, Summenformeln, physikalische Daten, IR- und Protonenresonanzspektren, chemische Reaktionen und deren Beschreibungen) sind direkt abfragbar gespeichert.
  Der Weg zu den Fakten über den Nachweis der Literaturzitate wird nicht gegangen.
- Zur Handhabung und Speicherung der Daten und Fakten werden Eingabeformulare, Speicherformate und Abfragecodes entwickelt und programmiert. Die Fragen wurden in Zusammenarbeit mit den Chemikern aus dem Arbeitskreis von Prof. Dr. W. Ried an der Universität Frankfurt erarbeitet.
- Zur Strukturierung der Daten und Fakten aus der Chlorformamidin-Chemie dienen algorithmische Vorschriften und heuristische Verfahren.

Heuristische Regeln greifen auf die physikalischen Eigenschaften und auf die Beschreibungen der Reaktionen zu und verarbeiten dieses Wissen zu neuen Fakten. Besonders ausgebaut wird die Heuristik, die sich mit den Reaktionen der Chlorformamidine und Wasser befaßt. Wasser dient bei der Untersuchung dieser Reaktion als Modellnukleophil, da es mit sehr vielen Chlorformamidinen glatt zu den gut zu charakterisierenden Harnstoffen reagiert. Die Reaktivitätsklassen der Chlorformamidine werden durch Abfragen nach physikalischen Eigenschaften und nach Daten zu Reaktionen gebildet.

In CHEMIS wurden etwa 25o Verbindungen und ebenso viele Reaktionen gespeichert, um die oben angesprochenen Konzepte zu realisieren.

Das System wurde von Herrn K. Ketterer auf der Großrechenanlage IBM /37o-3o31 der GID implementiert.

Durch die problemangepaßte Organisation des Wissens in Dateien entstand ein benutzerfreundliches Informationssystem für den Chemiker am Arbeitsplatz. Das System ist so aufgebaut, daß es wechselnden Anforderungen von Peptid-Chemikern und Chlorformamidin-Chemikern gerecht werden kann.

Leider konnte die Verarbeitung von Graphiken (Strukturformeln und Spektren) nicht realisiert werden, da die vorhandenen Bildschirme graphische Datenverarbeitung nicht zulassen.

8. Anhang

Alphabetisch geordnete Liste der Abfragecodes

Code	Erklärung
A,R_a	Ausbeuten der Reaktion a
A,V,R_a	Ausbeuten und Verbindungsnamen in der Reaktion a
EI,V_a	alle unter der Verbindungsnummer a gespeicherten Daten
FP,KP,V_a	Schmelz- und Siedepunkt einer Verbindung
G,R_a	alle Daten der Reaktionsgleicung a
INT	Liste der Interpretationen
IR,V_a	IR-Spektrum der Verbindung a
KAT,R_a	Katalysatoren der Reaktion a
KAT,V,R_a	Katalysatoren und Verbindungsnamen der Reaktion a
KP,V_a	Siedepunkt der Verbindung a
NMR,V_a	NMR-Spektrum der Verbindung a
L,R_a	Lösungsmittel in der Reaktion a
L,V,R_a	Lösungsmittel und Verbindungsnamen der Reaktion a

Code	Erklärung
P,R_a	Druck der Reaktion a
P,V,R_a	Druck und Verbindungsnamen der Reaktion a
R,KAT_a	alle Reaktionen mit dem Katalysator a
R,T_a	alle Reaktionsgleichungen mit der Reaktionstemperatur a
R,T_a_1; a_2	alle Reaktionsgleichungen, die im Temperaturintervall a_1 - a_2 ablaufen
RB,R_a	Reaktionsbedingungen der Gleichung a
RB,V,R_a	Reaktionsbedingungen und Stoffnamen der Gleichung a
SE	Liste der Schwingungselemente
SF,V_a	Summenformel der Verbindung a
V,CV_a	alle Verbindungen mit der chemischen Verschiebung a im NMR-Spektrum
V,CVN_a_1; ... a_{10}	alle Verbindungen mit den chemischen Verschiebungen a_1 bis a_{10} im NMR

Code	Erklärung
V,CVN,a_1; ... a_{10} (b)	alle Verbindungen mit den chemischen Verschiebungen a_1 bis a_{10}, wobei die Schwankungsbreite der Werte b betragen kann
V,FP_a	alle Verbingungen mit dem Schmelzpunkt a
V,FP_a_1; a_2	alle Verbindungen mit dem Schmelzpunkt im Intervall a_1 bis a_2
V,FSF_a	alle Verbindungen mit der Summenformel, die größer als a ist
V,INT,WZ_a,b	alle Verbindungen mit der Interpretation a und der Wellenzahl b im IR-Spektrum
V,INT,WZ_ a, b_1, b_2	alle Verbindungen mit der Interpretation a und dem Wellenzahlbereich b_1 bis b_2
V,IR,FP,INT_ a_1; a_2; b_1; ... b_6	alle Verbindungen mit einem Schmelzpunkt im Bereich a_1 bis a_2 und den Interpretationen b_1 bis b_6 im IR-Spektrum
V,IR,INT_a	alle Verbindungen, deren IR-Spektren Banden mit der Interprtation a enthalten

Code	Erklärung
V,KP_a	alle Verbindungen mit dem Siedepunkt a
V,KP_a_1, a_2	alle Verbindungen mit dem Siedepunktsbereich a_1 bis a_2
V,NMR,FP,INT_ a_1; a_2; b_1••• b_6	alle Verbindungen, die einen Schmelzpunkt im Bereich zwischen a_1 und a_2 haben und über die Interpretationen b_1 bis b_6 im NMR-Spektrum verfügen
V,NMR,KP,INT_ a_1; a_2; b_1••• b_6	alle Verbindungen, die einen Siedepunkt im Bereich a_1 bis a_2 haben und die über die Interpretation b_1 bis b_6 im NMR-Spektrum verfügen
V,NMR,INT_a	alle NMR-Spektren, die Peaks mit der Interpretation a aufweisen
V,R,KAT_a	alle Reaktionen und die Verbindungsnamen mit dem Katalysator a
V,R,T_a	alle Reaktionsgleichungen und die darin vorkommenden Verbindungsnamen, die bei der Temperatur a ablaufen
V,R,T_a_1, a_2	alle Reaktionsgleicungen und die darin vorkommenden Verbindungsnamen, die im Temperaturintervall von a_1 bis a_2 ablaufen
V,SF_a	alle Verbindungen mit der Summenformel a

Code	Erklärung
V,SE_a	alle Verbindungen mit dem Schwingungselement a
V,SE,WZ,a b_1,b_2	alle Verbindungen mit dem Schwingungselement a und dem Wellenzahlbereich b_1-b_2
V,TSF_a	alle Verbindungen mit der Summenformel, die a enthält
V,WZ_a	alle Verbindungen mit der Wellenzahl a im IR-Spektrum
V,WZN_a_1; a_2	alle Verbindungen mit dem Wellenzahlbereich a_1 bis a_2 im IR-Spektrum
V,WZN_a_1; ...a_{10}	alle Verbindungen mit den einzelnen Wellenzahlen a_1 bis a_{10}
V,WZN_a_1, ... a_{10}(b)	alle Verbindungen mit den Wellenzahlen a_1 bis a_{10},,Schwankungsbreite b
T,R_a	Reaktionstemperatur der Reaktion a
T,V,R_a	Reaktionstemperatur und Verbindungsnamen der Reaktion a

9. Literaturverzeichnis

(1) R. Capurro: "Information"
Verlag Dokumentation
München

(2) B. Rami: "Entwurf eines benutzerfreundlichen Informationssystems in der Peptidchemie"
Studiengruppe für Systemforschung
Bericht Nr. 121
Heidelberg 1976

(3) "CAS Today"
Chemical Abstracts Service
Columbus, Ohio 1980

(4) K. Arndt: "Wissenschaft, Informationskrise und Zeitschrift"
Chem.Techn. 25(12), S.745 (1973)

(5) D.J.de Solla-Price: "Science since Babylon"
Yale, 1961

(6) D.J.de Solla-Price: "Little Science, Big Science"
New York, London 1963

(7) C.E. Shannon, W. Weaver: "Mathematische Grundlagen der Informationstheorie"
München 1976

(8) N. Wiener: "Kybernetik"
Stuttgart 1965

(9) L. Brillouin: "Science and Information Theory"
New York 1962

(1o) G. Böhme: " Information und Verständigung"
in
E.v.Weizsäcker: "Offene Systeme I"
Stuttgart 1974

(11) C.F.v.Weizsäcker: "Die Einheit der Natur"
München 1979

(12) H. Borko: "Informationswissenschaft: Was ist das ?"
Nachr.Dok.19(3), S.59 (1968)

(13) H.-W. Schober, G.Wersig: "Informations- und Dokumentationswissenschaft"
Nachr.Dok. 19(4), S.116 (1968)

(14) G.Wersig: "Versuch zur Terminologie in der Dokumentation II: Kommunikation und Information"
Nachr.Dok. 2o(5), S.199 (1969)

(15) E.Pietsch: "Dokumentation und Information auf dem Weg zur Wissenschaft-Inhalt und Wandel der Begriffe"
Nachr.Dok. 19(6), S.199 (1968)

(16) W.Kunz: "Informationsverarbeitung in einer Erfahrungswissenschaft"
Nachr.Dok. 18(3/4), S. 85 (1967)

(17) H.Dietschmann, M.Ockenfeld, E.Pietsch: "Informationswissenschaften und Chemie
in
Kschenka et al. (Hrsg.) "Information und Dokumentation im Aufbruch"
München 1975

(18) W.L.Saunders: "The Nature of Information Science"
The Information Scientist, 8, S. 57 (1974)

(19) G.Hecht: "Grundlagen der Informationswissenschaft"
Informatik 22, S.5o (1975)

(2o) V.Slamecka: "Ein Forschungsprogramm für die Informationswissenschaft"
in
"3o Jahre DGD"
München 1979

(21) E.Garfield: "Where is Chemical Information Science Going ?"
J.Chem.Inf.Comp.Sci. 18(1), S.1 (1978)

(22) G.Wersig: "Information-Kommunikation-Dokumentation"
Darmstadt 1974

(23) A.I.Michajlov, A.I.Cernyj, R.S.Giljarevskij: "Wissenschaftliche Kommunikation und Informatik"
Leipzig 198o

(24) G.Kresze: "Kommunikation in der Chemie heute-Probleme und Möglichkeiten"
Angew.Chem. 82(15), S.563 (197o)

(25) R.Fugmann: "Theoretische Aspekte zur Kommunikation in der Chemie"
Angew.Chem. 82(15) S.574 (197o)

(26) G.M.Ahrenholz: "Dokumentation"
in
Ullmann's Enzyklopädie der technischen Chemie 4

(27) K.R.Popper: "Logik der Forschung"
Tübingen 1971

(28) K.R.Popper: "Objektive Erkenntnis"
Hamburg 1973

(29) Th.S.Kuhn: "Die Struktur wissenschaftlicher Revolutionen"
Frankfurt 1967

(3o) W.Kunz, H.Rittel: "Zur Logik von Forschung und Dokumentation"
Naturwissenschaften 1968, S.358

(31) H.J.Berliner: "A Chronology of Computer Chess and its Literature"
Artificial Intelligence 1o, S.2o1 (1978)

(32) D.Michie: "Chess with Computers"
Interdis.Sci.Rev. 5(3), S.215 (198o)

(33) J.W.Murdock: "Numerical Data Indexing"
J.Chem.Inf.Comput.Sci. 198o, S.132

(34) B.B. Molino: "Special Features of NBS Omnidata System Applicable to the Retrieval, Analysis,and Dissemination of Chemical Data"
J.Chem.Inf.Comput.Sci. 198o S.136

(35) R.C.Wilhoit: "The Design of a Multipurpose File of Thermodynamic Data"
J.Chem.Inf.Comput.Sci. 198o S.143

(36) D.T.Hawkins: "Problems in Physical Property Data Retrieval"
J.Chem.Inf.Comput.Sci. 1980 S.143

(37) G.C.Carter: "Numerical Data retrieval in the U.S. and Abroad"
J.Chem.Inf.Comput.Sci. 1980 S. 146

(38) D.M Kirchenbaum: "More Questions from a Data Compiler"
J.Chem.Inf.Comput.Sci. 1980 S.152

(39) "Kennen Sie Beilstein ?"
Frankfurt 1978

(40) R.Luckenbach, J.Sunkel: "Das wissenschaftliche Handbuch"
Naturwissenschaften 68(2) S.53 (1981)

(41) K.Chr.Buschbeck, W.Lippert, E.Ühlein: "Das systematische Handbuch in der naturwissenschaftlichen Literatur"
Naturwissenschaften 55(8) S.179 (1968)

(42) R.E.Burton, R.W.Kebler: "The "Half-Life" of some Scientific and Technical Literatures"
Amer.Doc. 11, S.18 (1960)

(43) P.Brown: "The Half-Life of the Chemical Literature"
J.Amer.Soc.Inf.Sci. 1980 S.61

(44) R.Walentowski: "Unique,Unambigous Representation of Chemical Structures by Computerization of a Simple Notation"
J.Chem.Inf.Comput.Sci. 1980, S.181

(45) "CAS Printed Access Tools"
CAS 1978

(46) E.Zass: "On-line-Recherchen in Datenbanken"
Naturwissenschaften, 69(12), 276, (1982)

(47) S.Lazerow: "Institut for Scientific Information"
in
"Encyclopedia of Library and Information Science"
A.Kent et al. (Ed.) 12, S.89
New York 1974

(48) M.Weinstock: "Citation Indexes"
in
"Encyclopedia of Library and Information Science"
A.Kent et al. (Ed.) 5, S.16
New York 1971

(49) H.Skolnik: "Historical Development of Abstracting"
J.Chem.Inf.Comput.Sci. 19(4), S.215 (1979)

(50) F.H.Barker: "Chemical Abstracts Material on-line"
Aslib Proc. 31(12), S.538 (1979)

(51) N.A.Farmer: "CAS online"
Database 1980, S.10

(52) H.Behrens,G.Ebel: "Data Compilations in Physics: Survey and Needs"
Proc. 6th CODATA Conf. S.221
Oxford 1979

(53) D.R.Lide: "Critical Data for Critical Needs"
Science,212(4501) S.1342 (1981)

(54) Mitteilung beim Beilstein-Workshop 1981

(55) M.A.Lobek: "Recherchen mit dem IDC-System"
Angew.Chem. 82(15) S.598 (1970)

(56) R.Fugmann, W.Braun, W.Vaupel
Angew.Chem. 73 S.745 (1961)

(57) J.E.Dubois et al.: "The DARC-System: A Chemical Information System Based on the Topological Encoding of Chemical Compounds"
Proc. Eurim II Conf. 1976 S. 135

(58) J.E Dubois: "DARC-System in Chemistry"
in
"Computer Representation and Manipulation of Chemical Information"
W.T.Wipke et al. (Ed.)
New York 1974

(59) J.Weizenbaum: "Die Macht der Computer und die Ohnmacht der Vernunft"
Frankfurt 1980

(60) B.G.Buchanan, E.A.Feigenbaum: "Dendral and Meta-Dendral"
Artificial Intelligence 11, S.5 (1978)

(61) J. Lederberg et al.: "Application of Artificial Intelligence for Chemical Inference I. The Number of Possible Organic Compounds: Acyclic Structures Containing C,H,O,and N"
J.Amer.Chem.Soc. 91 S.2973 (1969)

(62) E.C.Hass, P.J.Plath: "Graphentheoretische Modellierung der Katalyse von pericyclischen Reaktionen"
Match 7, S.299 (1979)

(63) E.J.Corey, W.T.Wipke: "Computer-Assisted Design of Complex Organic Syntheses"
Science 166, S.178 (1969)

(64) H.L.Gelernter et al.: "Empirical Explorations of SYNCHEM"
Science 197(4308), S.1041 (1977)

(65) M.Bersohn: "Automatic Problem Solving Applied to Synthetic Chemistry"
Bull.Chem.Soc.Jap. 45, S.1897 (1972)

(66) E.J.Corey: "Computer-Assisted Analysis of Complex Synthetic Problems"
Quart.Rev. (London) 25, S.455 (1971)

(67) J.Dugundji, I.Ugi: "An Algebraic Model of Constitutional Chemistry as a Basis for Chemical Computer Programs"
Top.Curr.Chem. 39 S.19 (1973)

(68) S.A.Rossmassler (Ed.): "Data Handling for Science and Technology"
Amsterdam 1980

(69) M.Bersohn, A.Esack: "Computers and Organic Syntheses"
Chem.Rev. 76(2), S.269 (1976)

(70) P.Gund et al.: "Computer-Assisted Synthetic Analysis at Merck"
J.Chem.Inf.Comput.Sci. 1980 S.88

(71) Einen Überblick bietet:
W.T.Wipke (Ed.): "Computer-Assisted Organic Syntheses"
American Chemical Society
Symposium Series Nr. 61 1977

(72) E.J.Corey et al.:"Computer-Assisted Synthetic Analysis for Complex Molecules. Methods and Procedures for Machine Generation of Synthetic Intermediates"
J.Amer.Chem.Soc. 94(2) S.44o (1972)

(73) E.J.Corey et al.: "Techniques for Perception by a Computer of Synthetically Significant Structural Features in Complex Molecules"
J.Amer.Chem.Soc. 94(2), S.431 (1972)

(74) E.J.Corey et al.: "Computer-Assisted Synthetic Analysis. Facile Man-Machine Communication of Chemical Structure by Interactive Computer-Graphics"
J.Amer.Chem.Soc. 94(2), S.421 (1972)

(75) W.Schubert, I.Ugi: "Darstellung chemischer Strukturen für die computergestützte deduktive Lösung chemischer Probleme"
Chimia 33(6), S.183 (1979)

(76) C.Jochum, J.Gasteiger, I.Ugi: "Das Prinzip der minimalen chemischen Distanz (PMCD)"
Angew.Chem. 92(7), S.5o3 (198o)

(77) J.F.Arens: "A Formalism for the Classification and Design of Organic Reactions I. The Class of $(+-)_n$ Reaktions"
Recl.Trav.Chim.Pays-Bas 98(4), S.155 (1979)

(78) J.F.Arens: "A Formalism for the Classification and Design of Organic Reaktions II. The Class of $(-+)_n$ Reaktions"
Recl.Trav.Chim.Pays-Bas 98(6), S.395 (1979)

(79) J.F.Arens: "A Formalism for the Classification and Design of Organic Reaktions III. The Class of $(+-)_nC$ Reaktions"
Recl.Trav.Chim.Pays-Bas 98(9),S.471 (1979)

(80) M.Ockenfeld: "Das Informationsverhalten von Chemikern"
München

(81) W.Kunz, H.Rittel: "Das Wissen des Chemikers und seine Information - Vorschläge zur informationswissenschaftlichen Ausbildung"
Chemiker Zeitung 96(1972), S.348

(82) W.Kunz, H.Rittel: "A Systems Analysis of the Logik of Research and Information Processes"
München 1977

(83) H. Ulrich: "The Chemistry of Imidoyl Halides"
New York 1968
Hier findet sich auch weitere Literatur zum Chlorformamidin-Gebiet.

(84) K.Bredereck, R.Richter: Ber. 99, S.2454 (1968)

(85) K.Bredereck, K.Richter: Ber. 99, S.2461 (1968)

(86) J.Fleming: "Selected Organic Syntheses"
London 1973

(87) Bei der Systemanalyse des Chlorformamidin-Gebiets wurde die folgende Literatur verwendet:
(87)-(113)
Reinhard Schweitzer: "Umlagerung und Cyclisierung N-Arylsubstituierter Amidin-Systeme"
Dissertation
Frankfurt 1976

(88) Hanns-Eberhard Erle: "Synthese und Reaktionen vinyloger Chlorformamidine"
Diplomarbeit
Frankfurt 1977

(89) Horst Winkler: "Synthese und Cyclisierung N-Aryl-substituierter Amidin-Systeme"
Diplomarbeit
Frankfurt 1977

(90) Hans Dietschmann: "Über Reaktionen von Halogencyclobutendionen mit Pseudohalogenen, Pseudohalogeniden und schwachen Stickstoffnukleophilen"
Dissertation
Frankfurt 1979

(91) Albrecht Höhle: "Über die Darstellung neuer Mono- und Bis-Chlorformamidine und einige ihrer Reaktionen"
Staatsexamensarbeit
Frankfurt 1976

(92) Rainer Christ: "Synthese und Reaktionen von und mit heterocyclischen Chlorformamidinen"
Dissertation
Frankfurt 1979

(93) Norbert Kothe: "Reaktionen mit Chlorformamidinen - Neue Wege für die Synthese von Chinazolinen"
Diplomarbeit
Frankfurt 1972

(94) Ulrike Vitt: "Studien über Reaktionen am Cyclobutendionsystem"
Diplomarbeit
Frankfurt 1979

(95) Norbert Kothe: "Umsetzungen mit Chlorformamidinen"
Dissertation
Frankfurt 1975

(96) Hanns-Eberhard Erle: "Über die Synthese neuartiger Chlorformamidine aus Dialkylcyanamiden und organischen Säurechloriden sowie deren Umsetzungen mit nukleophilen Reaktionspartnern"
Dissertation
Frankfurt 1981

(97) W.Ried, R.Guryn, G.Oremek: "Synthese und Reaktivität neuer heterocyclischer Chlorformamidine"
Chem.Ztg. 103(9), S.301 (1979)

(98) W.Ried, H.Dietschmann, H.-E.Erle: "Eine einfache Synthese von Carbamidoyl-halogeniden (Haloformamidinen) und von (dazu vinylogen) 2-Cyano-3,3-diaminopropenimidoyl-halogeniden"
Synthesis 198o(8), S.619

(99) W.Ried, G.Oremek, R.Guryn: "Notiz über die Synthese von Benzothiophenen"
Liebigs Ann. Chem. 196o, S.1172

(1oo) W.Ried, S.-W.Park: "Neue Heterocyclen aus substituierten 2,8-Dithia-4,1o-diaza-1,7-dimercurocyclododekanen"
J.f.prakt.Chem. 319(6), S.1o11 (1977)

(1o1) W.Ried, H.Knorr, G.Oremek: "Reaktionen von 4-Hydroxy--3-phenyl-cyclobut-(3)-en--onen-(1) mit Arylhydrazinen"
J.f.prakt.Chem. 3oo(2), S.246 (1978)

(1o2) H.Knorr, W.Ried, U.Knorr, G.Oremek: "Neue heterocyclische Systeme durch Umsetzung von 4-(-(z)-Benzyliden)-2,3-diphenyl-cyclobut-2-enon mit Diazomethan"
J.f.prakt.Chemie, 32o(4), S.573 (1978)

(1o3) W.Ried, G.Oremek, B.Ocakcioglu: "Synthese von substituierten Benzo(b)thiophenen"
Liebigs Ann.Chem. 198o S.1424

(1o4) W.Ried, G.Oremek, R.Guryn, H.-E. Erle: "Neuartige Oxazin-Synthese"
Chem.Ber. 113 S.2818 (198o)

(1o5) W.Ried, N.Kothe, R.Schweitzer, A.Höhle: "Synthese und Reaktionen bifunktioneller Chlorformamidine"
Chem.Ber. 1o9, S.2921 (1976)

(1o6) W.Ried, N.Kothe: "Über Umsetzungen von Chlorformamidinen und N-Phenyl-benzimidoylchlorid mit N-Cyanamidinen und 1-Cyanoguanidin"
Chem.Ber. 1o9, S.27o6 (1976)

(1o7) W.Ried, R.Schweitzer: "Neue Cyclisierungsprodukte aus N,N-diarylsubstituierter Propiolamidine"
Chem.Ber. 1o9, S.1643 (1976)

(1o8) W.Ried, W.Merkel, R.Schweitzer: "Über die Umlagerung von N-arylsubstituierten Nitryloxyformimidoyl--Verbindungen"
Chem.Ber. 1o8, S.229o (1975)

(1o9) W.Ried, W.Merkel: "Über die Umlagerung von N-arylsubstituierten Nitryloxyformamidinen"
Chem.Ber. 1o5, S.1532 (1972)

(11o) W.Ried, N.Kothe, W.Merkel: "Über Umsetzungen von Chlorformamidinen mit Cyanamiden"
Chem.Ber. 1o8, S.181 (1975)

(111) W.Ried, H.-E.Erle: "Synthesen und Reaktionen N-vinylsubstituierter Chlorformamidine"
Chem.Ber. 112, S.64o (1979)

(112) W.Ried, H.Winkler: "Cyclisierungsprodukte aus N-arylsubstituierten Propiolamidinen"
Chem.Ber. 112, S.384 (1979)

(98) W.Ried, H.Dietschmann, H.-E.Erle: "Eine einfache Synthese von Carbamidoyl-halogeniden (Haloformamidinen) und von (dazu vinylogen) 2-Cyano-3,3-diaminopropenimidoyl-halogeniden"
Synthesis 1980(8), S.619

(99) W.Ried, G.Oremek, R.Guryn: "Notiz über die Synthese von Benzothiophenen"
Liebigs Ann. Chem. 1980, S.1172

(100) W.Ried, S.-W.Park: "Neue Heterocyclen aus substituierten 2,8-Dithia-4,10-diaza-1,7-dimercurocyclododekanen"
J.f.prakt.Chem. 319(6), S.1011 (1977)

(101) W.Ried, H.Knorr, G.Oremek: "Reaktionen von 4-Hydroxy--3-phenyl-cyclobut-(3)-en--onen-(1) mit Arylhydrazinen"
J.f.prakt.Chem. 300(2), S.246 (1978)

(102) H.Knorr, W.Ried, U.Knorr, G.Oremek: "Neue heterocyclische Systeme durch Umsetzung von 4-(-(z)-Benzyliden)-2,3-diphenyl-cyclobut-2-enon mit Diazomethan"
J.f.prakt.Chemie, 320(4), S.573 (1978)

(103) W.Ried, G.Oremek, B.Ocakcioglu: "Synthese von substituierten Benzo(b)thiophenen"
Liebigs Ann.Chem. 1980 S.1424

(104) W.Ried, G.Oremek, R.Guryn, H.-E. Erle: "Neuartige Oxazin-Synthese"
Chem.Ber. 113 S.2818 (1980)

(1o5) W.Ried, N.Kothe, R.Schweitzer, A.Höhle: "Synthese und Reaktionen bifunktioneller Chlorformamidine"
Chem.Ber. 1o9, S.2921 (1976)

(1o6) W.Ried, N.Kothe: "Über Umsetzungen von Chlorformamidinen und N-Phenyl-benzimidoylchlorid mit N-Cyanamidinen und 1-Cyanoguanidin"
Chem.Ber. 1o9, S.27o6 (1976)

(1o7) W.Ried, R.Schweitzer: "Neue Cyclisierungsprodukte aus N,N-diarylsubstituierter Propiolamidine"
Chem.Ber. 1o9, S.1643 (1976)

(1o8) W.Ried, W.Merkel, R.Schweitzer: "Über die Umlagerung von N-arylsubstituierten Nitryloxyformimidoyl--Verbindungen"
Chem.Ber. 1o8, S.229o (1975)

(1o9) W.Ried, W.Merkel: "Über die Umlagerung von N-arylsubstituierten Nitryloxyformamidinen"
Chem.Ber. 1o5, S.1532 (1972)

(11o) W.Ried, N.Kothe, W.Merkel: "Über Umsetzungen von Chlorformamidinen mit Cyanamiden"
Chem.Ber. 1o8, S.181 (1975)

(111) W.Ried, H.-E.Erle: "Synthesen und Reaktionen N-vinylsubstituierter Chlorformamidine"
Chem.Ber. 112, S.64o (1979)

(112) W.Ried, H.Winkler: "Cyclisierungsprodukte aus N-arylsubstituierten Propiolamidinen"
Chem.Ber. 112, S.384 (1979)

(113) E.Kühle: "Kohlensäurederivate aus Formamiden"
Angew.Chem. 74(21), S.861 (1962)

(114) C.Wagner: "Methoden der naturwissenschaftlichen und technischen Forschung"
Mannheim 1974

(115) "Organikum"
Berlin 1972

(116) A,Nowak: "Fachliteratur des Chemikers"
Darmstadt 1976

(117) W.H.Littke, Ch.Jähn: "Rechnerische Identifikation in der infrarotspektroskopischen Analyse"
Z.Chem. 17(5), S.169 (1977)

(118) B.Franke, H.Pekar, H.Schweppe, H.Wagner: "Eine IR-Datenbank im täglichen Einsatz"
Fresenius Z.Anal.Chem. 303, S.349 (1980)

(119) J.-Th.Clerc, J.Zupan: "Computer-based Systems for the Retrieval of Infrared Spectral Data"
Pure&Appl. Chem. 49, S. 1827 (1977)

(120) Spektrensammlungen werden von
- Aldrich und
- Heyden & Sons (Sadtler-Spektren) herausgegeben.

(121) L.J.Bellamy: "Ultrarot-Spektrum und chemische Konstitution"
Darmstadt 1966

(122) O.F. Serebrjannikov: "Heuristische Prinzipien und logische Kalküle"
München-Salzburg 1974

(123) E.Oeser: "Wissenschaftstheorie als Rekonstruktion der Wissenschaftsgeschichte", Bd. 1
Wien-München 1979

(124) L.C.Smith: "Artificial Intelligence Applications in Information Systems"
in "Annual Review of Information Science and Technology" 15 (1980)
American Soc. Inf. Sci.

(125) V.A.Koptjug: "Über die Bedeutung von Spektrensammlungen und Rechenmaschinen in der chemischen Strukturforschung; Das wissenschaftliche Informationszentrum für Molekülspektroskopie in der Sibirischen Abteilung der Akademie der Wissenschaften der UdSSR"
Z.Chem. 15(2), S.41 (1975)

(126) F.Gerster, G.Dirlich: "Ein Programm zur Entscheidungshilfe in der Psychiatrie"
IBM Nachr. 25, S.224 (1975)

(127) J.Müller: "Operationen und Verfahren des problemlösenden Denkens in der konstruktiven technischen Entwicklungsarbeit - eine methodologische Studie"
Wiss. Z. d. Techn. Hochsch. Karl-Marx-Stadt 9(1/2)(1967)

(128) E.Oeser: "Wissenschaft und Information" Bd.I Wissenschaftstheorie und empirische Wissenschaftsforschung
Wien-München 1976

(129) I.Ugi, F.Beck, U.Fetzer: "Hydrolyse von Carbonsäureimidchloriden"
Chem.Ber. 95, S.126 (1962)

(130) A.F.Hegarty, J.D.Cronin, F.L.Scott: "Mechanism of Hydrolysis of Imidoyl Chlorides"
J.C.S. Perkin II 1975, S. 429

(131) R. Fugmann, J.H.Winter: "Durch mechanisierte Klassifikation zum Analogieschluß"
Vortrag auf der 3. Jahrestagung der Gesellschaft für Klassifikation Königstein, 1979

(132) H.G.O.Becker: "Einführung in die Elektronentheorie organisch-chemischer Reaktionen"
Berlin 1974

(133) CHEMIS Benutzer-Handbuch
Heidelberg, Dezember 1979

(134) I.Sebestyen: "The Videodisc Revolution"
Electronic Publishing Review, 2(1982) S.41

(135) J.H.Winter: "Chemische Syntheseplanung in Forschung und Industrie"
Heidelberg, 1982

(136) W.Kunz, H.Rittel: "Systemanalyse eiens Forschungs-
prozesses"
in: "Systemanalyse und Informationsverarbeitung in der Forschung" Hrsg. W.Kunz und H.Rittel
München 197o

(137) Computer-Assisted Drug Design
E.C.Olson, R.E.Christoffersen (Ed.)
ACS Symposium Series 112
Washington 1979

Mein Dank gilt:

- der Studiengruppe für Systemforschung und der Gesellschaft für Information und Dokumentation für die Förderung der Arbeit,

- Herrn K.Ketterer für die Programmierung des Systems,

- Frau Ockenfeld, Herrn Dietschmann und Herrn Nonnenmacher für die kritischen Diskussionen

und

- Herrn Prof.Dr.W.Ried für die Erlaubnis, die Chemie der Chlorformamidine, die von ihm und seinen Mitarbeitern untersucht wird, zu analysieren.

Meine akademischen Lehrer waren die
Professoren und Dozenten:

F.Becker, H.Bock, F.Comes, M.Czerny, K.Dänzer,
H.Fasold, H.Fritz, W.Grabner, H.Hartmann,
K.Helfrich, K.Hensen, H.Kessler, W.Kunz,
G.Kraft, W.Martienssen, E.Mutschler,
E.Pietsch, G.Quinkert, W.Ried, H.W.Roesky,
W.-D.Stohrer, D.Schaal, H.tomDiek, M.Trömel,
M.Wilk

Lebenslauf

21.10.1950	in Seligenstadt/Hessen geboren
Ostern 1956- Ostern 1960	Volksschule in Seligenstadt
Ostern 1960- Ostern 1969	Einhardschule-Gymnasium in Seligenstadt
01.07.1969- 30.06.1971	Bundeswehrdienstzeit
Wintersemester 1971	Beginn des Chemie-Studiums an der Johann-Wolfgang-Goethe-Universität in Frankfurt am Main
1973	Diplom-Chemiker Vorexamen
1975	Diplom-Chemiker Hauptexamen
1975- 1977	Anfertigung der Diplomarbeit im Arbeitskreis von Prof. Dr. W. Ried, Thema: "Über die Reaktionen von aromatischen Diaminen vom Typ des o-Phenylendiamins mit alpha-Acetylenketonen"
1977- 30.06.1980	Arbeit an der Dissertation als wissenschaftliche Hilfskraft bei der Studiengruppe für Systemforschung in Heidelberg, später Sektion für Systementwicklung der GID
ab 01.07.1980	wissenschaftlicher Referent bei der Gesellschaft für Information und Dokumentation im Zentralbereich Projektträgerschaft und Förderung in Frankfurt
30.12.1982	Heirat mit Irene Alice Jerat

Eidesstattliche Erklärung

Die vorliegende Arbeit wurde selbständig und nur unter Verwendung der zitierten Literatur und der angegebenen Hilfsmittel angefertigt.

Seligenstadt, Mai 1983

Eidesstattliche Erklärung

Die vorliegende Arbeit wurde selbständig und nur unter Verwendung der zitierten Literatur und der angegebenen Hilfsmittel angefertigt.

Seligenstadt, Mai 1983